

**UTILITY
PATENT APPLICATION
TRANSMITTAL**

Attorney Docket No.

210121.427C19

First Inventor or Application Identifier

Jiangchun Xu

Title

**COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF PROSTATE CANCER**

Express Mail Label No.

EL615232104US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO:Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 202311. ☐ General Authorization Form & Fee Transmittal
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages] **201**
(preferred arrangement set forth below)

- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention
- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claim(s)
- Abstract of the Disclosure

3. ☒ Drawing(s) (35 USC 113) [Total Sheets] **16**

4. Oath or Declaration [Total Pages]

a. ☐ Newly executed (original or copy)b. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)i. ☐ **DELETION OF INVENTOR(S)**
Signed statement attached deleting
inventor(s) named in the prior application,
see 37 CFR 1.63(d)(2) and 1.33(b)5. ☐ Incorporation By Reference (useable if box 4b is
checked) The entire disclosure of the prior application,
from which a copy of the oath or declaration is supplied
under Box 4b, is considered to be part of the disclosure of
the accompanying application and is hereby incorporated
by reference therein.6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and Amino Acid Sequence Submission
(if applicable, all necessary)a. ☒ Computer-Readable Copyb. ☒ Paper Copy (identical to computer copy)c. ☒ Statement verifying identity of above copies**ACCOMPANYING APPLICATION PARTS**8. ☐ Assignment Papers (cover sheet & document(s))9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney
(when there is an assignee)10. ☐ English Translation Document (if applicable)11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations12. ☐ Preliminary Amendment13. ☒ Return Receipt Postcard14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application,
Status still proper and desired15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)16. ☒ Other: Certificate of Express Mail

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information below and in a preliminary amendment

☐ Continuation ☐ Divisional ☒ Continuation-In-Part (CIP) of prior Application No.: not assignedPrior application information: Examiner not assignedGroup / Art Unit not assigned☐ Claims the benefit of Provisional Application No. _____**CORRESPONDENCE ADDRESS**Jane E. R. Potter
Seed Intellectual Property Law Group PLLC
701 Fifth Avenue, Suite 6300
Seattle, Washington 98104-7092
Phone: (206) 622-4900 Fax: (206) 682-6031

Respectfully submitted,

TYPED or PRINTED NAME Jane E. R. PotterSIGNATURE Jane E. R. Potter

u:\sharons\210121\427

REGISTRATION NO. 33,332Date September 6, 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Jiangchun Xu, Bellevue, WA; Davin C. Dillon, Issaquah, WA;
 Jennifer L. Mitcham, Redmond, WA; Susan L. Harlocker, Seattle,
 WA; Yuqiu Jiang, Kent, WA; Michael D. Kalos, Seattle, WA; Gary
 R. Fanger, Mill Creek, WA; Marc W. Retter, Carnation, WA; John A.
 Stolk, Bothell, WA; Craig H. Day, Seattle, WA; Thomas S. Vedvick,
 Federal Way, WA; Darrick Carter, Seattle, WA; Samuel X. Li,
 Redmond, WA; Aijun Wang, Issaquah, WA; Yasir A. W. Skeiky,
 Bellevue, WA; William Hepler, Seattle, WA; Robert A. Henderson,
 Edmonds, WA

Filed : September 6, 2000

For : COMPOSITIONS AND METHODS FOR THE THERAPY AND
 DIAGNOSIS OF PROSTATE CANCER

Docket No. : 210121.427C19

Date : September 6, 2000

Box Patent Application
 Assistant Commissioner for Patents
 Washington, DC 20231


CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Assistant Commissioner for Patents:

I hereby certify that the enclosures listed below are being deposited with the United States Postal Service "EXPRESS MAIL Post Office to Addressee" service under 37 C.F.R. § 1.10, Mailing Label Certificate No. EL615232104US, on September 6, 2000, addressed to Box Patent Application, Assistant Commissioner for Patents, Washington, DC 20231.

Respectfully submitted,

Seed Intellectual Property Law Group PLLC


 Steve Plante/Jeanette West/Susan Johnson

JEP:sds

Enclosures:

Postcard
 Form PTO/SB/05
 Specification, Claims, Abstract (201 pages)
 16 Sheets of Drawings (Figures 1-12)
 Sequence Listing (361 pages)
 Declaration for Sequence Listing
 Diskette for Sequence Listing

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF
PROSTATE CANCER

5 CROSS REFERENCE TO RELATED APPLICATIONS

This application is related to U.S. Patent Application No. 09/_____, filed August 29, 2000; U.S. Patent Application No. 09/636,215, filed August 9, 2000; U.S. Application No. 09/605,783, filed June 27, 2000; U.S. Patent Application No. 09/593,793, filed June 13, 2000; U.S. Patent Application No. 09/510,737, filed May 12, 2000; U.S. Patent Application No. 09/568,100, filed May 9, 2000; U.S. Patent Application No. 09/536,857, filed March 27, 2000; U. S. Patent Application No. 09/483,672, filed January 14, 2000; U.S. Patent Application No. 09/443,686, filed November 18, 1999; U.S. Patent Application No. 09/439,313, filed November 12, 1999; U.S. Patent Application No. 09/352,616, filed July 13, 1999; U.S. Patent Application No. 09/288,946, filed April 9, 1999; U.S. Patent Application No. 09/232,149, filed January 15, 1999; U.S. Patent Application No. 09/159,812, filed September 23, 1998; U.S. Patent Application No. 09/115,453, filed July 14, 1998; U.S. Patent Application No. 09/030,607, filed February 25, 1998; U.S. Patent Application No. 09/020,956, filed February 9, 1998; U.S. Patent Application No. 08/904,804, filed August 1, 1997; each a CIP of the previously filed application, and all pending, and U.S. Patent Application No. 08/806,099, filed February 25, 1997, now abandoned.

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as prostate cancer. The invention is more specifically related to polypeptides, comprising at least a portion of a prostate-specific protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides are useful in pharmaceutical

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

Prostate cancer is the most common form of cancer among males, with an estimated incidence of 30% in men over the age of 50. Overwhelming clinical evidence shows that human prostate cancer has the propensity to metastasize to bone, and the disease appears to progress inevitably from androgen dependent to androgen refractory status, leading to increased patient mortality. This prevalent disease is currently the second leading cause of cancer death among men in the U.S.

In spite of considerable research into therapies for the disease, prostate cancer remains difficult to treat. Commonly, treatment is based on surgery and/or radiation therapy, but these methods are ineffective in a significant percentage of cases. Two previously identified prostate specific proteins - prostate specific antigen (PSA) and prostatic acid phosphatase (PAP) - have limited therapeutic and diagnostic potential. For example, PSA levels do not always correlate well with the presence of prostate cancer, being positive in a percentage of non-prostate cancer cases, including benign prostatic hyperplasia (BPH). Furthermore, PSA measurements correlate with prostate volume, and do not indicate the level of metastasis.

In spite of considerable research into therapies for these and other cancers, prostate cancer remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

- (a) sequences provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;
- (b) complements of the sequences provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;
- (c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;
- (d) sequences that hybridize to a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, under moderately stringent conditions;
- (e) sequences having at least 75% identity to a sequence of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;
- (f) sequences having at least 90% identity to a sequence of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824; and
- (g) degenerate variants of a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and

384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824.

In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%, and
5 most preferably in at least about 50% of prostate tissue samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide sequence described
10 above.

The present invention further provides polypeptide compositions comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-
15 708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877.

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described
20 herein.

The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of
25 immunogenic activity of a polypeptide sequence set forth in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858 or 860-862, or a polypeptide sequence encoded by a polynucleotide sequence set forth in SEQ ID NO: 1-111, 115-171, 173-175,

177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, pharmaceutical compositions, *e.g.*, vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant, together with a physiologically acceptable carrier.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions, *e.g.*, vaccine

compositions, comprising a physiologically acceptable carrier and/or an immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating and/or enhancing the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with prostate cancer, in which case the methods provide treatment for the disease, or a patient considered to be at risk for such a disease may be treated prophylactically.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with prostate cancer, in which case the methods provide treatment for the disease, or a patient considered to be at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the polypeptide from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or

expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective
5 amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide
10 encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for
15 determining the presence or absence of a cancer, preferably a prostate cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a
20 cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a
25 binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b), and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide of the present invention; (b) detecting in the sample a level
 5 of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that
 10 hybridizes to a polynucleotide of the present invention, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to an inventive polynucleotide, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a
 15 cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide
 20 detected in step (c) with the amount detected in step (b), and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide
 25 probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

Figure 1 illustrates the ability of T cells to kill fibroblasts expressing the representative prostate-specific polypeptide P502S, as compared to control fibroblasts. The percentage lysis is shown as a series of effector:target ratios, as indicated.

Figure 3 represents a peptide competition binding assay showing that the P1S#10 peptide, derived from P501S, binds HLA-A2. Peptide P1S#10 inhibits HLA-A2 restricted presentation of fluM58 peptide to CTL clone D150M58 in TNF release bioassay. D150M58 CTL is specific for the HLA-A2 binding influenza matrix peptide fluM58.

Figure 5 illustrates the ability of a T cell clone to recognize and specifically
20 lyse Jurkat A2Kb cells expressing the representative prostate-specific polypeptide P501S,
thereby demonstrating that the P1S#10 peptide may be a naturally processed epitope of the
P501S polypeptide.

Figure 7 is a Western blot showing the expression of P501S in baculovirus.

Figure 8 illustrates the results of epitope mapping studies on P501S.

Figure 9 is a schematic representation of the P501S protein showing the location of transmembrane domains and predicted intracellular and extracellular domains.

Figure 10 is a genomic map showing the location of the prostate genes
 5 P775P, P704P, B305D, P712P and P774P within the Cat Eye Syndrome region of
 chromosome 22q11.2

Figure 11 shows the results of an ELISA assay to determine the specificity
 of rabbit polyclonal antisera raised against P501S.

Figures 12A(1), 12A(2), 12A(3), and B are the full-length cDNA (SEQ ID
 10 NO:591) and predicted amino acid (SEQ ID NO:592) sequences, respectively, for the clone
 P788P.

SEQ ID NO: 1 is the determined cDNA sequence for F1-13

SEQ ID NO: 2 is the determined 3' cDNA sequence for F1-12

SEQ ID NO: 3 is the determined 5' cDNA sequence for F1-12

15 SEQ ID NO: 4 is the determined 3' cDNA sequence for F1-16

SEQ ID NO: 5 is the determined 3' cDNA sequence for H1-1

SEQ ID NO: 6 is the determined 3' cDNA sequence for H1-9

SEQ ID NO: 7 is the determined 3' cDNA sequence for H1-4

SEQ ID NO: 8 is the determined 3' cDNA sequence for J1-17

20 SEQ ID NO: 9 is the determined 5' cDNA sequence for J1-17

SEQ ID NO: 10 is the determined 3' cDNA sequence for L1-12

SEQ ID NO: 11 is the determined 5' cDNA sequence for L1-12

SEQ ID NO: 12 is the determined 3' cDNA sequence for N1-1862

SEQ ID NO: 13 is the determined 5' cDNA sequence for N1-1862

25 SEQ ID NO: 14 is the determined 3' cDNA sequence for J1-13

SEQ ID NO: 15 is the determined 5' cDNA sequence for J1-13

SEQ ID NO: 16 is the determined 3' cDNA sequence for J1-19

SEQ ID NO: 17 is the determined 5' cDNA sequence for J1-19

SEQ ID NO: 18 is the determined 3' cDNA sequence for J1-25

SEQ ID NO: 19 is the determined 5' cDNA sequence for J1-25
 SEQ ID NO: 20 is the determined 5' cDNA sequence for J1-24
 SEQ ID NO: 21 is the determined 3' cDNA sequence for J1-24
 SEQ ID NO: 22 is the determined 5' cDNA sequence for K1-58
 5 SEQ ID NO: 23 is the determined 3' cDNA sequence for K1-58
 SEQ ID NO: 24 is the determined 5' cDNA sequence for K1-63
 SEQ ID NO: 25 is the determined 3' cDNA sequence for K1-63
 SEQ ID NO: 26 is the determined 5' cDNA sequence for L1-4
 SEQ ID NO: 27 is the determined 3' cDNA sequence for L1-4
 10 SEQ ID NO: 28 is the determined 5' cDNA sequence for L1-14
 SEQ ID NO: 29 is the determined 3' cDNA sequence for L1-14
 SEQ ID NO: 30 is the determined 3' cDNA sequence for J1-12
 SEQ ID NO: 31 is the determined 3' cDNA sequence for J1-16
 SEQ ID NO: 32 is the determined 3' cDNA sequence for J1-21
 15 SEQ ID NO: 33 is the determined 3' cDNA sequence for K1-48
 SEQ ID NO: 34 is the determined 3' cDNA sequence for K1-55
 SEQ ID NO: 35 is the determined 3' cDNA sequence for L1-2
 SEQ ID NO: 36 is the determined 3' cDNA sequence for L1-6
 SEQ ID NO: 37 is the determined 3' cDNA sequence for N1-1858
 20 SEQ ID NO: 38 is the determined 3' cDNA sequence for N1-1860
 SEQ ID NO: 39 is the determined 3' cDNA sequence for N1-1861
 SEQ ID NO: 40 is the determined 3' cDNA sequence for N1-1864
 SEQ ID NO: 41 is the determined cDNA sequence for P5
 SEQ ID NO: 42 is the determined cDNA sequence for P8
 25 SEQ ID NO: 43 is the determined cDNA sequence for P9
 SEQ ID NO: 44 is the determined cDNA sequence for P18
 SEQ ID NO: 45 is the determined cDNA sequence for P20
 SEQ ID NO: 46 is the determined cDNA sequence for P29
 SEQ ID NO: 47 is the determined cDNA sequence for P30

SEQ ID NO: 48 is the determined cDNA sequence for P34
 SEQ ID NO: 49 is the determined cDNA sequence for P36
 SEQ ID NO: 50 is the determined cDNA sequence for P38
 SEQ ID NO: 51 is the determined cDNA sequence for P39
 5 SEQ ID NO: 52 is the determined cDNA sequence for P42
 SEQ ID NO: 53 is the determined cDNA sequence for P47
 SEQ ID NO: 54 is the determined cDNA sequence for P49
 SEQ ID NO: 55 is the determined cDNA sequence for P50
 SEQ ID NO: 56 is the determined cDNA sequence for P53
 10 SEQ ID NO: 57 is the determined cDNA sequence for P55
 SEQ ID NO: 58 is the determined cDNA sequence for P60
 SEQ ID NO: 59 is the determined cDNA sequence for P64
 SEQ ID NO: 60 is the determined cDNA sequence for P65
 SEQ ID NO: 61 is the determined cDNA sequence for P73
 15 SEQ ID NO: 62 is the determined cDNA sequence for P75
 SEQ ID NO: 63 is the determined cDNA sequence for P76
 SEQ ID NO: 64 is the determined cDNA sequence for P79
 SEQ ID NO: 65 is the determined cDNA sequence for P84
 SEQ ID NO: 66 is the determined cDNA sequence for P68
 20 SEQ ID NO: 67 is the determined cDNA sequence for P80 (also referred to
 as P704P)
 SEQ ID NO: 68 is the determined cDNA sequence for P82
 SEQ ID NO: 69 is the determined cDNA sequence for U1-3064
 SEQ ID NO: 70 is the determined cDNA sequence for U1-3065
 25 SEQ ID NO: 71 is the determined cDNA sequence for V1-3692
 SEQ ID NO: 72 is the determined cDNA sequence for 1A-3905
 SEQ ID NO: 73 is the determined cDNA sequence for V1-3686
 SEQ ID NO: 74 is the determined cDNA sequence for R1-2330
 SEQ ID NO: 75 is the determined cDNA sequence for 1B-3976

SEQ ID NO: 76 is the determined cDNA sequence for V1-3679
 SEQ ID NO: 77 is the determined cDNA sequence for 1G-4736
 SEQ ID NO: 78 is the determined cDNA sequence for 1G-4738
 SEQ ID NO: 79 is the determined cDNA sequence for 1G-4741
 5 SEQ ID NO: 80 is the determined cDNA sequence for 1G-4744
 SEQ ID NO: 81 is the determined cDNA sequence for 1G-4734
 SEQ ID NO: 82 is the determined cDNA sequence for 1H-4774
 SEQ ID NO: 83 is the determined cDNA sequence for 1H-4781
 SEQ ID NO: 84 is the determined cDNA sequence for 1H-4785
 10 SEQ ID NO: 85 is the determined cDNA sequence for 1H-4787
 SEQ ID NO: 86 is the determined cDNA sequence for 1H-4796
 SEQ ID NO: 87 is the determined cDNA sequence for 1I-4807
 SEQ ID NO: 88 is the determined cDNA sequence for 1I-4810
 SEQ ID NO: 89 is the determined cDNA sequence for 1I-4811
 15 SEQ ID NO: 90 is the determined cDNA sequence for 1J-4876
 SEQ ID NO: 91 is the determined cDNA sequence for 1K-4884
 SEQ ID NO: 92 is the determined cDNA sequence for 1K-4896
 SEQ ID NO: 93 is the determined cDNA sequence for 1G-4761
 SEQ ID NO: 94 is the determined cDNA sequence for 1G-4762
 20 SEQ ID NO: 95 is the determined cDNA sequence for 1H-4766
 SEQ ID NO: 96 is the determined cDNA sequence for 1H-4770
 SEQ ID NO: 97 is the determined cDNA sequence for 1H-4771
 SEQ ID NO: 98 is the determined cDNA sequence for 1H-4772
 SEQ ID NO: 99 is the determined cDNA sequence for 1D-4297
 25 SEQ ID NO: 100 is the determined cDNA sequence for 1D-4309
 SEQ ID NO: 101 is the determined cDNA sequence for 1D.1-4278
 SEQ ID NO: 102 is the determined cDNA sequence for 1D-4288
 SEQ ID NO: 103 is the determined cDNA sequence for 1D-4283
 SEQ ID NO: 104 is the determined cDNA sequence for 1D-4304

SEQ ID NO: 105 is the determined cDNA sequence for 1D-4296

SEQ ID NO: 106 is the determined cDNA sequence for 1D-4280

SEQ ID NO: 107 is the determined full length cDNA sequence for F1-12
(also referred to as P504S)

5 SEQ ID NO: 108 is the predicted amino acid sequence for F1-12

SEQ ID NO: 109 is the determined full length cDNA sequence for J1-17

SEQ ID NO: 110 is the determined full length cDNA sequence for L1-12
(also referred to as P501S)

10 SEQ ID NO: 111 is the determined full length cDNA sequence for N1-1862
(also referred to as P503S)

SEQ ID NO: 112 is the predicted amino acid sequence for J1-17

SEQ ID NO: 113 is the predicted amino acid sequence for L1-12 (also
referred to as P501S)

15 SEQ ID NO: 114 is the predicted amino acid sequence for N1-1862 (also
referred to as P503S)

SEQ ID NO: 115 is the determined cDNA sequence for P89

SEQ ID NO: 116 is the determined cDNA sequence for P90

SEQ ID NO: 117 is the determined cDNA sequence for P92

SEQ ID NO: 118 is the determined cDNA sequence for P95

20 SEQ ID NO: 119 is the determined cDNA sequence for P98

SEQ ID NO: 120 is the determined cDNA sequence for P102

SEQ ID NO: 121 is the determined cDNA sequence for P110

SEQ ID NO: 122 is the determined cDNA sequence for P111

SEQ ID NO: 123 is the determined cDNA sequence for P114

25 SEQ ID NO: 124 is the determined cDNA sequence for P115

SEQ ID NO: 125 is the determined cDNA sequence for P116

SEQ ID NO: 126 is the determined cDNA sequence for P124

SEQ ID NO: 127 is the determined cDNA sequence for P126

SEQ ID NO: 128 is the determined cDNA sequence for P130

SEQ ID NO: 129 is the determined cDNA sequence for P133
 SEQ ID NO: 130 is the determined cDNA sequence for P138
 SEQ ID NO: 131 is the determined cDNA sequence for P143
 SEQ ID NO: 132 is the determined cDNA sequence for P151
 5 SEQ ID NO: 133 is the determined cDNA sequence for P156
 SEQ ID NO: 134 is the determined cDNA sequence for P157
 SEQ ID NO: 135 is the determined cDNA sequence for P166
 SEQ ID NO: 136 is the determined cDNA sequence for P176
 SEQ ID NO: 137 is the determined cDNA sequence for P178
 10 SEQ ID NO: 138 is the determined cDNA sequence for P179
 SEQ ID NO: 139 is the determined cDNA sequence for P185
 SEQ ID NO: 140 is the determined cDNA sequence for P192
 SEQ ID NO: 141 is the determined cDNA sequence for P201
 SEQ ID NO: 142 is the determined cDNA sequence for P204
 15 SEQ ID NO: 143 is the determined cDNA sequence for P208
 SEQ ID NO: 144 is the determined cDNA sequence for P211
 SEQ ID NO: 145 is the determined cDNA sequence for P213
 SEQ ID NO: 146 is the determined cDNA sequence for P219
 SEQ ID NO: 147 is the determined cDNA sequence for P237
 20 SEQ ID NO: 148 is the determined cDNA sequence for P239
 SEQ ID NO: 149 is the determined cDNA sequence for P248
 SEQ ID NO: 150 is the determined cDNA sequence for P251
 SEQ ID NO: 151 is the determined cDNA sequence for P255
 SEQ ID NO: 152 is the determined cDNA sequence for P256
 25 SEQ ID NO: 153 is the determined cDNA sequence for P259
 SEQ ID NO: 154 is the determined cDNA sequence for P260
 SEQ ID NO: 155 is the determined cDNA sequence for P263
 SEQ ID NO: 156 is the determined cDNA sequence for P264
 SEQ ID NO: 157 is the determined cDNA sequence for P266

SEQ ID NO: 158 is the determined cDNA sequence for P270
 SEQ ID NO: 159 is the determined cDNA sequence for P272
 SEQ ID NO: 160 is the determined cDNA sequence for P278
 SEQ ID NO: 161 is the determined cDNA sequence for P105
 5 SEQ ID NO: 162 is the determined cDNA sequence for P107
 SEQ ID NO: 163 is the determined cDNA sequence for P137
 SEQ ID NO: 164 is the determined cDNA sequence for P194
 SEQ ID NO: 165 is the determined cDNA sequence for P195
 SEQ ID NO: 166 is the determined cDNA sequence for P196
 10 SEQ ID NO: 167 is the determined cDNA sequence for P220
 SEQ ID NO: 168 is the determined cDNA sequence for P234
 SEQ ID NO: 169 is the determined cDNA sequence for P235
 SEQ ID NO: 170 is the determined cDNA sequence for P243
 SEQ ID NO: 171 is the determined cDNA sequence for P703P-DE1
 15 SEQ ID NO: 172 is the predicted amino acid sequence for P703P-DE1
 SEQ ID NO: 173 is the determined cDNA sequence for P703P-DE2
 SEQ ID NO: 174 is the determined cDNA sequence for P703P-DE6
 SEQ ID NO: 175 is the determined cDNA sequence for P703P-DE13
 SEQ ID NO: 176 is the predicted amino acid sequence for P703P-DE13
 20 SEQ ID NO: 177 is the determined cDNA sequence for P703P-DE14
 SEQ ID NO: 178 is the predicted amino acid sequence for P703P-DE14
 SEQ ID NO: 179 is the determined extended cDNA sequence for 1G-4736
 SEQ ID NO: 180 is the determined extended cDNA sequence for 1G-4738
 SEQ ID NO: 181 is the determined extended cDNA sequence for 1G-4741
 25 SEQ ID NO: 182 is the determined extended cDNA sequence for 1G-4744
 SEQ ID NO: 183 is the determined extended cDNA sequence for 1H-4774
 SEQ ID NO: 184 is the determined extended cDNA sequence for 1H-4781
 SEQ ID NO: 185 is the determined extended cDNA sequence for 1H-4785
 SEQ ID NO: 186 is the determined extended cDNA sequence for 1H-4787

SEQ ID NO: 187 is the determined extended cDNA sequence for 1H-4796
 SEQ ID NO: 188 is the determined extended cDNA sequence for 1I-4807
 SEQ ID NO: 189 is the determined 3' cDNA sequence for 1I-4810
 SEQ ID NO: 190 is the determined 3' cDNA sequence for 1I-4811
 5 SEQ ID NO: 191 is the determined extended cDNA sequence for 1J-4876
 SEQ ID NO: 192 is the determined extended cDNA sequence for 1K-4884
 SEQ ID NO: 193 is the determined extended cDNA sequence for 1K-4896
 SEQ ID NO: 194 is the determined extended cDNA sequence for 1G-4761
 SEQ ID NO: 195 is the determined extended cDNA sequence for 1G-4762
 10 SEQ ID NO: 196 is the determined extended cDNA sequence for 1H-4766
 SEQ ID NO: 197 is the determined 3' cDNA sequence for 1H-4770
 SEQ ID NO: 198 is the determined 3' cDNA sequence for 1H-4771
 SEQ ID NO: 199 is the determined extended cDNA sequence for 1H-4772
 SEQ ID NO: 200 is the determined extended cDNA sequence for 1D-4309
 15 SEQ ID NO: 201 is the determined extended cDNA sequence for 1D.1-4278
 SEQ ID NO: 202 is the determined extended cDNA sequence for 1D-4288
 SEQ ID NO: 203 is the determined extended cDNA sequence for 1D-4283
 SEQ ID NO: 204 is the determined extended cDNA sequence for 1D-4304
 SEQ ID NO: 205 is the determined extended cDNA sequence for 1D-4296
 20 SEQ ID NO: 206 is the determined extended cDNA sequence for 1D-4280
 SEQ ID NO: 207 is the determined cDNA sequence for 10-d8fwd
 SEQ ID NO: 208 is the determined cDNA sequence for 10-H10con
 SEQ ID NO: 209 is the determined cDNA sequence for 11-C8rev
 SEQ ID NO: 210 is the determined cDNA sequence for 7.g6fwd
 25 SEQ ID NO: 211 is the determined cDNA sequence for 7.g6rev
 SEQ ID NO: 212 is the determined cDNA sequence for 8-b5fwd
 SEQ ID NO: 213 is the determined cDNA sequence for 8-b5rev
 SEQ ID NO: 214 is the determined cDNA sequence for 8-b6fwd
 SEQ ID NO: 215 is the determined cDNA sequence for 8-b6 rev

SEQ ID NO: 216 is the determined cDNA sequence for 8-d4fwd
 SEQ ID NO: 217 is the determined cDNA sequence for 8-d9rev
 SEQ ID NO: 218 is the determined cDNA sequence for 8-g3fwd
 SEQ ID NO: 219 is the determined cDNA sequence for 8-g3rev
 5 SEQ ID NO: 220 is the determined cDNA sequence for 8-h11rev
 SEQ ID NO: 221 is the determined cDNA sequence for g-fl2fwd
 SEQ ID NO: 222 is the determined cDNA sequence for g-f3rev
 SEQ ID NO: 223 is the determined cDNA sequence for P509S
 SEQ ID NO: 224 is the determined cDNA sequence for P510S
 10 SEQ ID NO: 225 is the determined cDNA sequence for P703DE5
 SEQ ID NO: 226 is the determined cDNA sequence for 9-A11
 SEQ ID NO: 227 is the determined cDNA sequence for 8-C6
 SEQ ID NO: 228 is the determined cDNA sequence for 8-H7
 SEQ ID NO: 229 is the determined cDNA sequence for JPTPN13
 15 SEQ ID NO: 230 is the determined cDNA sequence for JPTPN14
 SEQ ID NO: 231 is the determined cDNA sequence for JPTPN23
 SEQ ID NO: 232 is the determined cDNA sequence for JPTPN24
 SEQ ID NO: 233 is the determined cDNA sequence for JPTPN25
 SEQ ID NO: 234 is the determined cDNA sequence for JPTPN30
 20 SEQ ID NO: 235 is the determined cDNA sequence for JPTPN34
 SEQ ID NO: 236 is the determined cDNA sequence for PTPN35
 SEQ ID NO: 237 is the determined cDNA sequence for JPTPN36
 SEQ ID NO: 238 is the determined cDNA sequence for JPTPN38
 SEQ ID NO: 239 is the determined cDNA sequence for JPTPN39
 25 SEQ ID NO: 240 is the determined cDNA sequence for JPTPN40
 SEQ ID NO: 241 is the determined cDNA sequence for JPTPN41
 SEQ ID NO: 242 is the determined cDNA sequence for JPTPN42
 SEQ ID NO: 243 is the determined cDNA sequence for JPTPN45
 SEQ ID NO: 244 is the determined cDNA sequence for JPTPN46

SEQ ID NO: 245 is the determined cDNA sequence for JPTPN51
 SEQ ID NO: 246 is the determined cDNA sequence for JPTPN56
 SEQ ID NO: 247 is the determined cDNA sequence for PTPN64
 SEQ ID NO: 248 is the determined cDNA sequence for JPTPN65
 5 SEQ ID NO: 249 is the determined cDNA sequence for JPTPN67
 SEQ ID NO: 250 is the determined cDNA sequence for JPTPN76
 SEQ ID NO: 251 is the determined cDNA sequence for JPTPN84
 SEQ ID NO: 252 is the determined cDNA sequence for JPTPN85
 SEQ ID NO: 253 is the determined cDNA sequence for JPTPN86
 10 SEQ ID NO: 254 is the determined cDNA sequence for JPTPN87
 SEQ ID NO: 255 is the determined cDNA sequence for JPTPN88
 SEQ ID NO: 256 is the determined cDNA sequence for JP1F1
 SEQ ID NO: 257 is the determined cDNA sequence for JP1F2
 SEQ ID NO: 258 is the determined cDNA sequence for JP1C2
 15 SEQ ID NO: 259 is the determined cDNA sequence for JP1B1
 SEQ ID NO: 260 is the determined cDNA sequence for JP1B2
 SEQ ID NO: 261 is the determined cDNA sequence for JP1D3
 SEQ ID NO: 262 is the determined cDNA sequence for JP1A4
 SEQ ID NO: 263 is the determined cDNA sequence for JP1F5
 20 SEQ ID NO: 264 is the determined cDNA sequence for JP1E6
 SEQ ID NO: 265 is the determined cDNA sequence for JP1D6
 SEQ ID NO: 266 is the determined cDNA sequence for JP1B5
 SEQ ID NO: 267 is the determined cDNA sequence for JP1A6
 SEQ ID NO: 268 is the determined cDNA sequence for JP1E8
 25 SEQ ID NO: 269 is the determined cDNA sequence for JP1D7
 SEQ ID NO: 270 is the determined cDNA sequence for JP1D9
 SEQ ID NO: 271 is the determined cDNA sequence for JP1C10
 SEQ ID NO: 272 is the determined cDNA sequence for JP1A9
 SEQ ID NO: 273 is the determined cDNA sequence for JP1F12

SEQ ID NO: 274 is the determined cDNA sequence for JP1E12
 SEQ ID NO: 275 is the determined cDNA sequence for JP1D11
 SEQ ID NO: 276 is the determined cDNA sequence for JP1C11
 SEQ ID NO: 277 is the determined cDNA sequence for JP1C12
 5 SEQ ID NO: 278 is the determined cDNA sequence for JP1B12
 SEQ ID NO: 279 is the determined cDNA sequence for JP1A12
 SEQ ID NO: 280 is the determined cDNA sequence for JP8G2
 SEQ ID NO: 281 is the determined cDNA sequence for JP8H1
 SEQ ID NO: 282 is the determined cDNA sequence for JP8H2
 10 SEQ ID NO: 283 is the determined cDNA sequence for JP8A3
 SEQ ID NO: 284 is the determined cDNA sequence for JP8A4
 SEQ ID NO: 285 is the determined cDNA sequence for JP8C3
 SEQ ID NO: 286 is the determined cDNA sequence for JP8G4
 SEQ ID NO: 287 is the determined cDNA sequence for JP8B6
 15 SEQ ID NO: 288 is the determined cDNA sequence for JP8D6
 SEQ ID NO: 289 is the determined cDNA sequence for JP8F5
 SEQ ID NO: 290 is the determined cDNA sequence for JP8A8
 SEQ ID NO: 291 is the determined cDNA sequence for JP8C7
 SEQ ID NO: 292 is the determined cDNA sequence for JP8D7
 20 SEQ ID NO: 293 is the determined cDNA sequence for P8D8
 SEQ ID NO: 294 is the determined cDNA sequence for JP8E7
 SEQ ID NO: 295 is the determined cDNA sequence for JP8F8
 SEQ ID NO: 296 is the determined cDNA sequence for JP8G8
 SEQ ID NO: 297 is the determined cDNA sequence for JP8B10
 25 SEQ ID NO: 298 is the determined cDNA sequence for JP8C10
 SEQ ID NO: 299 is the determined cDNA sequence for JP8E9
 SEQ ID NO: 300 is the determined cDNA sequence for JP8E10
 SEQ ID NO: 301 is the determined cDNA sequence for JP8F9
 SEQ ID NO: 302 is the determined cDNA sequence for JP8H9

- SEQ ID NO: 303 is the determined cDNA sequence for JP8C12
 SEQ ID NO: 304 is the determined cDNA sequence for JP8E11
 SEQ ID NO: 305 is the determined cDNA sequence for JP8E12
 SEQ ID NO: 306 is the amino acid sequence for the peptide PS2#12
 5 SEQ ID NO: 307 is the determined cDNA sequence for P711P
 SEQ ID NO: 308 is the determined cDNA sequence for P712P
 SEQ ID NO: 309 is the determined cDNA sequence for CLONE23
 SEQ ID NO: 310 is the determined cDNA sequence for P774P
 SEQ ID NO: 311 is the determined cDNA sequence for P775P
 10 SEQ ID NO: 312 is the determined cDNA sequence for P715P
 SEQ ID NO: 313 is the determined cDNA sequence for P710P
 SEQ ID NO: 314 is the determined cDNA sequence for P767P
 SEQ ID NO: 315 is the determined cDNA sequence for P768P
 SEQ ID NO: 316-325 are the determined cDNA sequences of previously
 15 isolated genes
 SEQ ID NO: 326 is the determined cDNA sequence for P703PDE5
 SEQ ID NO: 327 is the predicted amino acid sequence for P703PDE5
 SEQ ID NO: 328 is the determined cDNA sequence for P703P6.26
 SEQ ID NO: 329 is the predicted amino acid sequence for P703P6.26
 20 SEQ ID NO: 330 is the determined cDNA sequence for P703PX-23
 SEQ ID NO: 331 is the predicted amino acid sequence for P703PX-23
 SEQ ID NO: 332 is the determined full length cDNA sequence for P509S
 SEQ ID NO: 333 is the determined extended cDNA sequence for P707P
 (also referred to as 11-C9)
 25 SEQ ID NO: 334 is the determined cDNA sequence for P714P
 SEQ ID NO: 335 is the determined cDNA sequence for P705P (also
 referred to as 9-F3)
 SEQ ID NO: 336 is the predicted amino acid sequence for P705P
 SEQ ID NO: 337 is the amino acid sequence of the peptide P1S#10

- SEQ ID NO: 338 is the amino acid sequence of the peptide p5
- SEQ ID NO: 339 is the predicted amino acid sequence of P509S
- SEQ ID NO: 340 is the determined cDNA sequence for P778P
- SEQ ID NO: 341 is the determined cDNA sequence for P786P
- 5 SEQ ID NO: 342 is the determined cDNA sequence for P789P
- SEQ ID NO: 343 is the determined cDNA sequence for a clone showing
homology to Homo sapiens MM46 mRNA
- SEQ ID NO: 344 is the determined cDNA sequence for a clone showing
homology to Homo sapiens TNF-alpha stimulated ABC protein (ABC50) mRNA
- 10 SEQ ID NO: 345 is the determined cDNA sequence for a clone showing
homology to Homo sapiens mRNA for E-cadherin
- SEQ ID NO: 346 is the determined cDNA sequence for a clone showing
homology to Human nuclear-encoded mitochondrial serine hydroxymethyltransferase
(SHMT)
- 15 SEQ ID NO: 347 is the determined cDNA sequence for a clone showing
homology to Homo sapiens natural resistance-associated macrophage protein2 (NRAMP2)
- SEQ ID NO: 348 is the determined cDNA sequence for a clone showing
homology to Homo sapiens phosphoglucomutase-related protein (PGMRP)
- SEQ ID NO: 349 is the determined cDNA sequence for a clone showing
20 homology to Human mRNA for proteosome subunit p40
- SEQ ID NO: 350 is the determined cDNA sequence for P777P
- SEQ ID NO: 351 is the determined cDNA sequence for P779P
- SEQ ID NO: 352 is the determined cDNA sequence for P790P
- SEQ ID NO: 353 is the determined cDNA sequence for P784P
- 25 SEQ ID NO: 354 is the determined cDNA sequence for P776P
- SEQ ID NO: 355 is the determined cDNA sequence for P780P
- SEQ ID NO: 356 is the determined cDNA sequence for P544S
- SEQ ID NO: 357 is the determined cDNA sequence for P745S
- SEQ ID NO: 358 is the determined cDNA sequence for P782P

- SEQ ID NO: 359 is the determined cDNA sequence for P783P
- SEQ ID NO: 360 is the determined cDNA sequence for unknown 17984
- SEQ ID NO: 361 is the determined cDNA sequence for P787P
- SEQ ID NO: 362 is the determined cDNA sequence for P788P
- 5 SEQ ID NO: 363 is the determined cDNA sequence for unknown 17994
- SEQ ID NO: 364 is the determined cDNA sequence for P781P
- SEQ ID NO: 365 is the determined cDNA sequence for P785P
- SEQ ID NO: 366-375 are the determined cDNA sequences for splice variants of B305D.
- 10 SEQ ID NO: 376 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 366.
- SEQ ID NO: 377 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 372.
- SEQ ID NO: 378 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 373.
- 15 SEQ ID NO: 379 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 374.
- SEQ ID NO: 380 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 375.
- 20 SEQ ID NO: 381 is the determined cDNA sequence for B716P.
- SEQ ID NO: 382 is the determined full-length cDNA sequence for P711P.
- SEQ ID NO: 383 is the predicted amino acid sequence for P711P.
- SEQ ID NO: 384 is the cDNA sequence for P1000C.
- SEQ ID NO: 385 is the cDNA sequence for CGI-82.
- 25 SEQ ID NO:386 is the cDNA sequence for 23320.
- SEQ ID NO:387 is the cDNA sequence for CGI-69.
- SEQ ID NO:388 is the cDNA sequence for L-iditol-2-dehydrogenase.
- SEQ ID NO:389 is the cDNA sequence for 23379.
- SEQ ID NO:390 is the cDNA sequence for 23381.

SEQ ID NO:391 is the cDNA sequence for KIAA0122.

SEQ ID NO:392 is the cDNA sequence for 23399.

SEQ ID NO:393 is the cDNA sequence for a previously identified gene.

SEQ ID NO:394 is the cDNA sequence for HCLBP.

5 SEQ ID NO:395 is the cDNA sequence for transglutaminase.

SEQ ID NO:396 is the cDNA sequence for a previously identified gene.

SEQ ID NO:397 is the cDNA sequence for PAP.

SEQ ID NO:398 is the cDNA sequence for Ets transcription factor PDEF.

SEQ ID NO:399 is the cDNA sequence for hTGR.

10 SEQ ID NO:400 is the cDNA sequence for KIAA0295.

SEQ ID NO:401 is the cDNA sequence for 22545.

SEQ ID NO:402 is the cDNA sequence for 22547.

SEQ ID NO:403 is the cDNA sequence for 22548.

SEQ ID NO:404 is the cDNA sequence for 22550.

15 SEQ ID NO:405 is the cDNA sequence for 22551.

SEQ ID NO:406 is the cDNA sequence for 22552.

SEQ ID NO:407 is the cDNA sequence for 22553 (also known as P1020C).

SEQ ID NO:408 is the cDNA sequence for 22558.

SEQ ID NO:409 is the cDNA sequence for 22562.

20 SEQ ID NO:410 is the cDNA sequence for 22565.

SEQ ID NO:411 is the cDNA sequence for 22567.

SEQ ID NO:412 is the cDNA sequence for 22568.

SEQ ID NO:413 is the cDNA sequence for 22570.

SEQ ID NO:414 is the cDNA sequence for 22571.

25 SEQ ID NO:415 is the cDNA sequence for 22572.

SEQ ID NO:416 is the cDNA sequence for 22573.

SEQ ID NO:417 is the cDNA sequence for 22573.

SEQ ID NO:418 is the cDNA sequence for 22575.

SEQ ID NO:419 is the cDNA sequence for 22580.

SEQ ID NO:420 is the cDNA sequence for 22581.
 SEQ ID NO:421 is the cDNA sequence for 22582.
 SEQ ID NO:422 is the cDNA sequence for 22583.
 SEQ ID NO:423 is the cDNA sequence for 22584.
 5 SEQ ID NO:424 is the cDNA sequence for 22585.
 SEQ ID NO:425 is the cDNA sequence for 22586.
 SEQ ID NO:426 is the cDNA sequence for 22587.
 SEQ ID NO:427 is the cDNA sequence for 22588.
 SEQ ID NO:428 is the cDNA sequence for 22589.
 10 SEQ ID NO:429 is the cDNA sequence for 22590.
 SEQ ID NO:430 is the cDNA sequence for 22591.
 SEQ ID NO:431 is the cDNA sequence for 22592.
 SEQ ID NO:432 is the cDNA sequence for 22593.
 SEQ ID NO:433 is the cDNA sequence for 22594.
 15 SEQ ID NO:434 is the cDNA sequence for 22595.
 SEQ ID NO:435 is the cDNA sequence for 22596.
 SEQ ID NO:436 is the cDNA sequence for 22847.
 SEQ ID NO:437 is the cDNA sequence for 22848.
 SEQ ID NO:438 is the cDNA sequence for 22849.
 20 SEQ ID NO:439 is the cDNA sequence for 22851.
 SEQ ID NO:440 is the cDNA sequence for 22852.
 SEQ ID NO:441 is the cDNA sequence for 22853.
 SEQ ID NO:442 is the cDNA sequence for 22854.
 SEQ ID NO:443 is the cDNA sequence for 22855.
 25 SEQ ID NO:444 is the cDNA sequence for 22856.
 SEQ ID NO:445 is the cDNA sequence for 22857.
 SEQ ID NO:446 is the cDNA sequence for 23601.
 SEQ ID NO:447 is the cDNA sequence for 23602.
 SEQ ID NO:448 is the cDNA sequence for 23605.

SEQ ID NO:449 is the cDNA sequence for 23606.
 SEQ ID NO:450 is the cDNA sequence for 23612.
 SEQ ID NO:451 is the cDNA sequence for 23614.
 SEQ ID NO:452 is the cDNA sequence for 23618.
 5 SEQ ID NO:453 is the cDNA sequence for 23622.
 SEQ ID NO:454 is the cDNA sequence for folate hydrolase.
 SEQ ID NO:455 is the cDNA sequence for LIM protein.
 SEQ ID NO:456 is the cDNA sequence for a known gene.
 SEQ ID NO:457 is the cDNA sequence for a known gene.
 10 SEQ ID NO:458 is the cDNA sequence for a previously identified gene.
 SEQ ID NO:459 is the cDNA sequence for 23045.
 SEQ ID NO:460 is the cDNA sequence for 23032.
 SEQ ID NO:461 is the cDNA sequence for 23054.
 SEQ ID NO:462-467 are cDNA sequences for known genes.
 15 SEQ ID NO:468-471 are cDNA sequences for P710P.
 SEQ ID NO:472 is a cDNA sequence for P1001C.
 SEQ ID NO: 473 is the determined cDNA sequence for a first splice variant
 of P775P (referred to as 27505).
 SEQ ID NO: 474 is the determined cDNA sequence for a second splice
 20 variant of P775P (referred to as 19947).
 SEQ ID NO: 475 is the determined cDNA sequence for a third splice variant
 of P775P (referred to as 19941).
 SEQ ID NO: 476 is the determined cDNA sequence for a fourth splice
 variant of P775P (referred to as 19937).
 25 SEQ ID NO: 477 is a first predicted amino acid sequence encoded by the
 sequence of SEQ ID NO: 474.
 SEQ ID NO: 478 is a second predicted amino acid sequence encoded by the
 sequence of SEQ ID NO: 474.

SEQ ID NO: 523 is a mature form of P703P used to raise antibodies against P703P.

SEQ ID NO: 524 is the putative full-length cDNA sequence of P703P.

SEQ ID NO: 525 is the predicted amino acid sequence encoded by SEQ ID
5 NO: 524.

SEQ ID NO: 526 is the full-length cDNA sequence for P790P.

SEQ ID NO: 527 is the predicted amino acid sequence for P790P.

SEQ ID NO: 528 & 529 are PCR primers.

SEQ ID NO: 530 is the cDNA sequence of a splice variant of SEQ ID NO:
10 366.

SEQ ID NO: 531 is the cDNA sequence of the open reading frame of SEQ
ID NO: 530.

SEQ ID NO: 532 is the predicted amino acid encoded by the sequence of
SEQ ID NO: 531.

15 SEQ ID NO: 533 is the DNA sequence of a putative ORF of P775P.

SEQ ID NO: 534 is the predicted amino acid sequence encoded by SEQ ID
NO: 533.

SEQ ID NO: 535 is a first full-length cDNA sequence for P510S.

SEQ ID NO: 536 is a second full-length cDNA sequence for P510S.

20 SEQ ID NO: 537 is the predicted amino acid sequence encoded by SEQ ID
NO: 535.

SEQ ID NO: 538 is the predicted amino acid sequence encoded by SEQ ID
NO: 536.

SEQ ID NO: 539 is the peptide P501S-370.

25 SEQ ID NO: 540 is the peptide P501S-376.

SEQ ID NO: 541-551 are epitopes of P501S.

SEQ ID NO: 552 is an extended cDNA sequence for P712P.

SEQ ID NO: 553-568 are the amino acid sequences encoded by predicted
open reading frames within SEQ ID NO: 552.

SEQ ID NO: 570 is the determined cDNA sequence for a splice variant of P776P referred to as contig 6.

SEQ ID NO: 571 is the determined cDNA sequence for a splice variant of P776P referred to as contig 7.

SEQ ID NO: 572 is the determined cDNA sequence for a splice variant of P776P referred to as contig 14.

SEQ ID NO: 573 is the amino acid sequence encoded by a first predicted ORF of SEQ ID NO: 570.

SEQ ID NO: 574 is the amino acid sequence encoded by a second predicted ORF of SEQ ID NO: 570.

SEQ ID NO: 575 is the amino acid sequence encoded by a predicted ORF of SEQ ID NO: 571.

SEQ ID NO: 576-586 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 569.

SEQ ID NO: 587 is a DNA consensus sequence of the sequences of P767P and P777P.

SEQ ID NO: 588-590 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 587.

SEQ ID NO: 591 is an extended cDNA sequence for P1020C.

SEQ ID NO: 592 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: P1020C.

SEQ ID NO: 593 is a splice variant of P775P referred to as 50748.

SEQ ID NO: 594 is a splice variant of P775P referred to as 50717.

SEQ ID NO: 595 is a splice variant of P775P referred to as 45985.

SEQ ID NO: 596 is a splice variant of P775P referred to as 38769.

SEQ ID NO: 597 is a splice variant of P775P referred to as 37922.

SEQ ID NO: 598 is a splice variant of P510S referred to as 49274.

SEQ ID NO: 599 is a splice variant of P510S referred to as 39487.

SEQ ID NO: 600 is a splice variant of P504S referred to as 5167.16.
 SEQ ID NO: 601 is a splice variant of P504S referred to as 5167.1.
 SEQ ID NO: 602 is a splice variant of P504S referred to as 5163.46.
 SEQ ID NO: 603 is a splice variant of P504S referred to as 5163.42.
 5 SEQ ID NO: 604 is a splice variant of P504S referred to as 5163.34.
 SEQ ID NO: 605 is a splice variant of P504S referred to as 5163.17.
 SEQ ID NO: 606 is a splice variant of P501S referred to as 10640.
 SEQ ID NO: 607-615 are the sequences of PCR primers.
 SEQ ID NO: 616 is the determined cDNA sequence of a fusion of P703P
 10 and PSA.
 SEQ ID NO: 617 is the amino acid sequence of the fusion of P703P and
 PSA.
 SEQ ID NO: 618-689 are determined cDNA sequences of prostate-specific
 clones.
 15 SEQ ID NO: 690 is the cDNA sequence of the gene DD3.
 SEQ ID NO: 691-697 are determined cDNA sequences of prostate-specific
 clones.
 SEQ ID NO: 698 is an extended cDNA sequence for P714P.
 SEQ ID NO: 699-701 are the cDNA sequences for splice variants of P704P.
 20 SEQ ID NO: 702 is the cDNA sequence of a spliced variant of P553S
 referred to as P553S-14.
 SEQ ID NO: 703 is the cDNA sequence of a spliced variant of P553S
 referred to as P553S-12.
 SEQ ID NO: 704 is the cDNA sequence of a spliced variant of P553S
 25 referred to as P553S-10.
 SEQ ID NO: 705 is the cDNA sequence of a spliced variant of P553S
 referred to as P553S-6.
 SEQ ID NO: 706 is the amino acid sequence encoded by SEQ ID NO: 705.

SEQ ID NO: 707 is the amino acid sequence encoded by SEQ ID NO: 702
SEQ ID NO: 708 is a second amino acid sequence encoded by SEQ ID NO: 702.

SEQ ID NO: 709-772 are determined cDNA sequences of prostate-specific clones.

5 SEQ ID NO: 773 is a first full-length cDNA sequence for prostate-specific transglutaminase gene (also referred to herein as P558S).

SEQ ID NO: 774 is a second full-length cDNA sequence for prostate-specific transglutaminase gene.

10 SEQ ID NO: 775 is the amino acid sequence encoded by the sequence of SEQ ID NO: 773.

SEQ ID NO: 776 is the amino acid sequence encoded by the sequence of SEQ ID NO: 774.

SEQ ID NO: 777 is the full-length cDNA sequence for P788P.

SEQ ID NO: 778 is the amino acid sequence encoded by SEQ ID NO: 777.

15 SEQ ID NO: 779 is the determined cDNA sequence for a polymorphic variant of P788P.

SEQ ID NO: 780 is the amino acid sequence encoded by SEQ ID NO: 779.

SEQ ID NO: 781 is the amino acid sequence of peptide 4 from P703P.

SEQ ID NO: 782 is the cDNA sequence that encodes peptide 4 from P703P.

20 SEQ ID NO: 783-798 are the cDNA sequence encoding epitopes of P703P.

SEQ ID NO: 799-814 are the amino acid sequences of epitopes of P703P.

SEQ ID NO: 815 and 816 are PCR primers.

SEQ ID NO: 817 is the cDNA sequence encoding an N-terminal portion of P788P expressed in *E. coli*.

25 SEQ ID NO: 818 is the amino acid sequence of the N-terminal portion of P788P expressed in *E. coli*.

SEQ ID NO: 819 is the amino acid sequence of the *M. tuberculosis* antigen Ra12.

SEQ ID NO: 820 and 821 are PCR primers.

- SEQ ID NO: 822 is the cDNA sequence for the Ra12-P510S-C construct.
- SEQ ID NO: 823 is the cDNA sequence for the P510S-C construct.
- SEQ ID NO: 824 is the cDNA sequence for the P510S-E3 construct.
- SEQ ID NO: 825 is the amino acid sequence for the Ra12-P510S-C
 5 construct.
- SEQ ID NO: 826 is the amino acid sequence for the P510S-C construct.
- SEQ ID NO: 827 is the amino acid sequence for the P510S-E3 construct.
- SEQ ID NO: 828-833 are PCR primers.
- SEQ ID NO: 834 is the cDNA sequence of the construct Ra12-P775P-
 10 ORF3.
- SEQ ID NO: 835 is the amino acid sequence of the construct Ra12-P775P-
 ORF3.
- SEQ ID NO: 836 and 837 are PCR primers.
- SEQ ID NO: 838 is the determined amino acid sequence for a P703P His tag
 15 fusion protein.
- SEQ ID NO: 839 is the determined cDNA sequence for a P703P His tag
 fusion protein.
- SEQ ID NO: 840 and 841 are PCR primers.
- SEQ ID NO: 842 is the determined amino acid sequence for a P705P His tag
 20 fusion protein.
- SEQ ID NO: 843 is the determined cDNA sequence for a P705P His tag
 fusion protein.
- SEQ ID NO: 844 and 845 are PCR primers.
- SEQ ID NO: 846 is the determined amino acid sequence for a P711P His tag
 25 fusion protein.
- SEQ ID NO: 847 is the determined cDNA sequence for a P711P His tag
 fusion protein.
- SEQ ID NO: 848 is the amino acid sequence of the *M. tuberculosis* antigen
 Ra12.

SEQ ID NO: 849 and 850 are PCR primers.

SEQ ID NO: 851 is the determined cDNA sequence for the construct Ra12-P501S-E2.

5 SEQ ID NO: 852 is the determined amino acid sequence for the construct Ra12-P501S-E2.

SEQ ID NO: 853 is the amino acid sequence for an epitope of P501S.

SEQ ID NO: 854 is the DNA sequence encoding SEQ ID NO: 853.

SEQ ID NO: 855 is the amino acid sequence for an epitope of P501S.

SEQ ID NO: 856 is the DNA sequence encoding SEQ ID NO: 855.

10 SEQ ID NO: 857 is a peptide employed in epitope mapping studies.

SEQ ID NO: 858 is the amino acid sequence for an epitope of P501S.

SEQ ID NO: 859 is the DNA sequence encoding SEQ ID NO: 858.

SEQ ID NO: 860-862 are the amino acid sequences for CD4 epitopes of P501S.

15 SEQ ID NO: 863-865 are the DNA sequences encoding the sequences of SEQ ID NO: 860-862.

SEQ ID NO: 866-877 are the amino acid sequences for putative CTL epitopes of P703P.

20 DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly prostate cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (*e.g.*, T cells).

25

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology,

molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, et al. *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis et al. *Molecular Cloning: A Laboratory Manual* (1982);
 5 DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984).

All publications, patents and patent applications cited herein, whether supra
 10 or infra, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

Polypeptide Compositions

As used herein, the term "polypeptide" " is used in its conventional
 15 meaning, *i.e.*, as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations,
 20 phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, *i.e.*, antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking
 25 an immune response.

Particularly illustrative polypeptides of the present invention comprise those encoded by a polynucleotide sequence set forth in any one of SEQ ID NOs: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-

476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NOs: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824. In specific embodiments, the polypeptides of the invention comprise amino acid sequences as set forth in any one of SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877.

The polypeptides of the present invention are sometimes herein referred to as prostate-specific proteins or prostate-specific polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in prostate tissue samples. Thus, a "prostate-specific polypeptide" or "prostate-specific protein," refers generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial proportion of prostate tissue samples, for example preferably greater than about 20%, more preferably greater than about 30%, and most preferably greater than about 50% or more of prostate tissue samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in other normal tissues, as determined using a representative assay provided herein. A prostate-specific polypeptide sequence of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described below.

In certain preferred embodiments, the polypeptides of the invention are immunogenic, *i.e.*, they react detectably within an immunoassay (such as an ELISA or T-cell stimulation assay) with antisera and/or T-cells from a patient with prostate cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those

described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies
 5 detected using, for example, ^{125}I -labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An “immunogenic portion,” as used herein, is a fragment of an immunogenic polypeptide of the invention that itself is immunologically reactive (*i.e.*, specifically binds) with the B-
 10 cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used
 15 herein, antisera and antibodies are “antigen-specific” if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not
 20 substantially less than the reactivity of the full-length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length polypeptide. In some instances,
 25 preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length polypeptide, *e.g.*, having greater than about 100% or 150% or more immunogenic activity.

In certain other embodiments, illustrative immunogenic portions may include peptides in which an N-terminal leader sequence and/or transmembrane domain has

been deleted. Other illustrative immunogenic portions will contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide composition set forth herein, such as those set forth in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877, or those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequence set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provided by the present invention are immunologically reactive with an antibody and/or T-
5 cell that reacts with a full-length polypeptide specifically set forth herein.

In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or more of that exhibited by a full-length polypeptide sequence specifically set forth herein.

10 A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as
15 described herein using any of a number of techniques well known in the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been
20 removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially
25 unchanged. As described above, modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or

portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes,

substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8);
 5 glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with
 10 similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101
 15 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0);
 20 threonine (−0.4); proline (−0.5 \pm 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically
 25 equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-, methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other

sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be
 5 “identical” if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to
 10 about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment
 15 schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183,
 20 Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J.
 25 (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl.*

Acad. Sci. USA 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining
 5 percent sequence identity and sequence similarity are the BLAST and BLAST 2.0
 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402
 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0
 can be used, for example with the parameters described herein, to determine percent
 sequence identity for the polynucleotides and polypeptides of the invention. Software for
 10 performing BLAST analyses is publicly available through the National Center for
 Biotechnology Information. For amino acid sequences, a scoring matrix can be used to
 calculate the cumulative score. Extension of the word hits in each direction are halted
 when: the cumulative alignment score falls off by the quantity X from its maximum
 achieved value; the cumulative score goes to zero or below, due to the accumulation of one
 15 or more negative-scoring residue alignments; or the end of either sequence is reached. The
 BLAST algorithm parameters W, T and X determine the sensitivity and speed of the
 alignment.

In one preferred approach, the "percentage of sequence identity" is
 determined by comparing two optimally aligned sequences over a window of comparison
 20 of at least 20 positions, wherein the portion of the polypeptide sequence in the comparison
 window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to
 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not
 comprise additions or deletions) for optimal alignment of the two sequences. The
 percentage is calculated by determining the number of positions at which the identical
 25 amino acid residue occurs in both sequences to yield the number of matched positions,
 dividing the number of matched positions by the total number of positions in the reference
 sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage
 of sequence identity.

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes.

Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 5 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or 10 translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The fusion polypeptide can comprise a polypeptide as described herein 15 together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, 336:86-91, 1997).

In one preferred embodiment, the immunological fusion partner is derived from a *Mycobacterium* sp., such as a *Mycobacterium tuberculosis*-derived Ra12 fragment. 20 Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 25 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; *see also*, Skeiky et al., *Infection and Immun.* (1999) 67:3998-4007, incorporated herein by reference). C-terminal fragments of the MTB32A coding sequence express at high levels and remain as a soluble

polypeptides throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid residues 192 to 323 of MTB32A. Other preferred Ra12 polynucleotides

5 generally comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide. Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a

10 sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a

15 polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D

20 derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural

25 protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known

as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of CD4⁺ T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See Merrifield, J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original

environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, *e.g.*, are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

5 Polynucleotide Compositions

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a
 10 polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

15 As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

20 As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or
 25 non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide sequence set forth in any one of SEQ ID NOs: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, complements of a polynucleotide sequence set forth in any one of SEQ ID NOs: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, and degenerate variants of a polynucleotide sequence set forth in any one of SEQ ID NOs: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824. In certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

50

nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompass homologous genes of xenogenic origin.

In additional embodiments, the present invention provides polynucleotide fragments comprising various lengths of contiguous stretches of sequence identical to, or complementary to, one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the

hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, *e.g.*,
 5 polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth
 10 herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length
 15 may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the
 20 like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be “identical” if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are
 25 typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, preferably 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of
 5 evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS*
 10 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

15 Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP,
 20 BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402
 25 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for

nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides.

The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an

increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term “oligonucleotide directed mutagenesis procedure” is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to “evolve” individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 contiguous nucleotides that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned,

such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced

using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by which a disease can be treated by inhibiting the synthesis of proteins that contribute to the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, Science. 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein. Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and

determination of secondary structure, T_m , binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly preferred target regions of the mRNA, are those which are at
 5 or near the AUG translation initiation codon, and those sequences which are substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997 Sep 1;25(17):3389-402).

10 The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several molecules of
 15 the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

 According to another embodiment of the invention, the polynucleotide
 20 compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, Proc Natl Acad Sci U S A. 1987 Dec;84(24):8788-
 25 92; Forster and Symons, Cell. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, Cell. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, J Mol Biol. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, Nature. 1992 May 14;357(6374):173-6).

This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7305-9). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of

5 hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.*, Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis δ virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada *et al.*,

10 Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is

15 that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO

20 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

25 Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688,

which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

5 Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and
10 bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation,
15 oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within
20 cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene
25 regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells. Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but

not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, *Antisense Nucleic Acid Drug Dev.* 1997 7(4) 431-37). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997 Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Hanvey *et al.*, *Science*. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, *Bioorg Med Chem.* 1996 Jan;4(1):5-23). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, *Bioorg Med Chem.* 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of

chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45; Petersen *et al.*, J Pept Sci. 1995 May-Jun;1(3):175-83; Orum *et al.*, Biotechniques. 1995 Sep;19(3):472-80; Footer *et al.*, Biochemistry. 1996 Aug 20;35(33):10673-9; Griffith *et al.*, Nucleic Acids Res. 1995 Aug 11;23(15):3003-8; Pardridge *et al.*, Proc Natl Acad Sci U S A. 1995 Jun 6;92(12):5592-6; Boffa *et al.*, Proc Natl Acad Sci U S A. 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, Blood. 1996 Aug 15;88(4):1411-7; Armitage *et al.*, Proc Natl Acad Sci U S A. 1997 Nov 11;94(23):12320-5; Seeger *et al.*, Biotechniques. 1997 Sep;23(3):512-7). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (Anal Chem. 1993 Dec 15;65(24):3545-9) and Jensen *et al.*

(Biochemistry. 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

5 Other applications of PNAs that have been described and will be apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition, mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

10 Polynucleotide Identification, Characterization and Expression

Polynucleotide compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a
 15 polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's
 20 instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target
 25 sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are

complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended
 5 along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification procedure may be performed in order to quantify the amount of mRNA amplified.
 10 Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCRTM amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent No.
 15 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No.
 20 WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl. Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence
 25 to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, a tumor cDNA library) using

well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred
 5 for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see*
 10 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences
 15 may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

20 Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template
 25 for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure,

which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to
 5 identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by
 10 analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

15 In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally
 20 equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or
 25 eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding

sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-
5 directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion
10 protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

15 Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example,
20 peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable
25 techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered

during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression
 5 vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in*
 10 *vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and
 15 express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus,
 20 TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription
 25 and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPO11 plasmid (Gibco BRL,

Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable
 5 marker.

In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors
 10 include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and
 15 the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease
 20 cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.*
 25 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987)

EMBO J. 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG

initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the

introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art.

These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include, for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

5 A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to
10 two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

15 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be
20 cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides,
25 enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained

intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

Antibody Compositions, Fragments Thereof and Other Binding Agents

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunologically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off}/K_{on} enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant K_d . See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or

"FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as prostate cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of

monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the

yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and
 5 extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves
 10 IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "F(ab')₂" fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA
 15 immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent V_H::V_L heterodimer including an antigen-binding site which retains much of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710;
 20 and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked V_H::V_L heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical
 25 structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRS and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (*e.g.*, a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRS. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human

constant domains (Winter et al. (1991) Nature 349:293-299; Lobuglio et al. (1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw et al. (1987) J Immunol. 138:4534-4538; and Brown et al. (1987) Cancer Res. 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain
 5 (Riechmann et al. (1988) Nature 332:323-327; Verhoeyen et al. (1988) Science 239:1534-1536; and Jones et al. (1986) Nature 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the
 10 duration and effectiveness of therapeutic applications of those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, *e.g.*, a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule
 15 comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) Ann. Rev. Biochem. 59:439-473. Thus, antigen
 20 binding specificity can be preserved in a humanized antibody only wherein the CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior (*e.g.*, solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises
 25 either a weakly immunogenic, or substantially non-immunogenic veneered surface.

The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in Sequences of Proteins of Immunological Interest, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible

U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest
 5 are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V regions are then compared residue by residue to corresponding murine amino acids. The residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only
 10 carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "veneered" murine antigen-binding sites are
 15 thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (*e.g.*, electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR
 20 loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

25 In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred

differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile

bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

5 It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled
10 directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides
15 such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and
20 their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

T Cell Compositions

25 The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or

peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide,

polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

5 For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the
10 addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

Pharmaceutical Compositions

15 In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may
20 be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the
25 particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide, antibody, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and therapeutic vaccine applications. Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, *e.g.*, vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (*e.g.*, U.S. Pat. No. 5,219,740; Miller and Rosman (1989) *BioTechniques* 7:980-990; Miller, A. D. (1990) *Human Gene Therapy* 1:5-14; Scarpa et al. (1991) *Virology* 180:849-852; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033-8037; and Boris-Lawrie and Temin (1993) *Cur. Opin. Genet. Develop.* 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) *J. Virol.* 57:267-274; Bett et al. (1993) *J. Virol.* 67:5911-5921; Mittereder et al. (1994) *Human Gene Therapy* 5:717-729; Seth et al. (1994) *J. Virol.* 68:933-940; Barr et al. (1994) *Gene Therapy* 1:51-58; Berkner, K. L. (1988) *BioTechniques* 6:616-629; and Rich et al. (1993) *Human Gene Therapy* 4:461-476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, *e.g.*, U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al. (1988) *Molec. Cell. Biol.* 8:3988-3996; Vincent et al. (1990) *Vaccines* 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) *Current Opinion in Biotechnology* 3:533-539; Muzyczka, N. (1992) *Current Topics in Microbiol. and Immunol.* 158:97-129; Kotin, R. M. (1994) *Human Gene Therapy* 5:793-801; Shelling and Smith (1994) *Gene Therapy* 1:165-169; and Zhou et al. (1994) *J. Exp. Med.* 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The

5 DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome.

10 The resulting TK.sup.(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first

15 infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected

20 DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, *e.g.*, Elroy-Stein and Moss, *Proc. Natl. Acad. Sci. USA* (1990) 87:6743-6747; Fuerst et al. *Proc. Natl. Acad. Sci. USA* (1986) 83:8122-8126.

25 Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox

genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

5 Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

10 Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al. *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based
 15 delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991;
 20 Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993.

In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in a specific location and orientation *via*
 25 homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner

in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum

hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL[®] adjuvants are available from Corixa Corporation (Seattle, WA; *see*, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555,

WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or
 5 *Gypsophila* or *Chenopodium quinoa* saponins. Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β -escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine
 10 vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or
 15 ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamellar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol^R to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

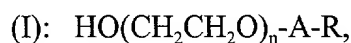
20 In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL[®] adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another
 25 particularly preferred adjuvant formulation employing QS21, 3D-MPL[®] adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG

and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhanzyn[®]; Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula



wherein, n is 1-50, A is a bond or $-\text{C}(\text{O})-$, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein n is between 1 and 50, preferably 4-24, most preferably 9; the R component is C_{1-50} , preferably $\text{C}_4\text{-C}_{20}$ alkyl and most preferably C_{12} alkyl, and A is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12th edition: entry 7717). These adjuvant molecules are described in WO 99/52549. The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium

combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc γ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for
5 example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments,
10 however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which
15 comprise a non-liquid hydrophilic core (e.g., a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and
20 expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (e.g., polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344,
25 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems, such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents
5 such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-
10 dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately
15 prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of
20 illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or
25 they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998;15(3):243-84; U.

S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration, the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water,

binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about

by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art.

Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not
 5 produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described, *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent
 10 5,804,212. Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045.

15 In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions
 20 of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998 Mar;56(3):691-
 25 5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, the use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

Cancer Therapeutic Methods

In further aspects of the present invention, the pharmaceutical compositions described herein may be used for the treatment of cancer, particularly for the immunotherapy of prostate cancer. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably

a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use

intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever et al., *Immunological Reviews* 157:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the

basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

Cancer Detection and Diagnostic Compositions, Methods and Kits

In general, a cancer may be detected in a patient based on the presence of one or more prostate tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as prostate cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a

prostate tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length prostate tumor proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent

No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which
5 may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1
10 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be
15 achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*,
20 Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound
25 sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with prostate cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may

generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as prostate cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the

sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μ g/ml). It may be desirable to

incubate another aliquot of a T cell sample in the absence of tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for

example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further,

5 The present invention further provides kits for use within any of the above
diagnostic methods. Such kits typically comprise two or more components necessary for
performing a diagnostic assay. Components may be compounds, reagents, containers
and/or equipment. For example, one container within a kit may contain a monoclonal
antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or
10 fragments may be provided attached to a support material, as described above. One or
more additional containers may enclose elements, such as reagents or buffers, to be used in
the assay. Such kits may also, or alternatively, contain a detection reagent as described
above that contains a reporter group suitable for direct or indirect detection of antibody
binding.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES

This Example describes the isolation of certain prostate-specific
5 polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library was constructed from
prostate tumor poly A⁺ RNA using a Superscript Plasmid System for cDNA Synthesis and
Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the
manufacturer's protocol. Specifically, prostate tumor tissues were homogenized with
10 polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent
(BRL Life Technologies) as directed by the manufacturer. The poly A⁺ RNA was then
purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa
Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was
synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized,
15 ligated with EcoRI/BAXI adaptors (Invitrogen, San Diego, CA) and digested with NotI.
Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA),
the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen) and transformed
into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

Using the same procedure, a normal human pancreas cDNA expression
20 library was prepared from a pool of six tissue specimens (Clontech). The cDNA libraries
were characterized by determining the number of independent colonies, the percentage of
clones that carried insert, the average insert size and by sequence analysis. The prostate
tumor library contained 1.64×10^7 independent colonies, with 70% of clones having an
insert and the average insert size being 1745 base pairs. The normal pancreas cDNA
25 library contained 3.3×10^6 independent colonies, with 69% of clones having inserts and the
average insert size being 1120 base pairs. For both libraries, sequence analysis showed that
the majority of clones had a full length cDNA sequence and were synthesized from mRNA,
with minimal rRNA and mitochondrial DNA contamination.

cDNA library subtraction was performed using the above prostate tumor and normal pancreas cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a prostate tumor-specific subtracted cDNA library was generated as follows. Normal pancreas cDNA library (70 µg) was digested with
 5 EcoRI, NotI, and SfuI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 µl of H₂O, heat-denatured and mixed with 100 µl (100 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional
 10 Photoprobe biotin (50 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H₂O to form the driver DNA.

To form the tracer DNA, 10 µg prostate tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400
 15 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H₂O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]). The reaction
 20 mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H₂O, mixed with 8 µl driver DNA and 20 µl of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into
 25 BamHI/XhoI site of chloramphenicol resistant pBCSK⁺ (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a prostate tumor specific subtracted cDNA library (referred to as "prostate subtraction 1").

To analyze the subtracted cDNA library, plasmid DNA was prepared from 100 independent clones, randomly picked from the subtracted prostate tumor specific

library and grouped based on insert size. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA). Six cDNA clones, hereinafter referred to as F1-13, F1-12, F1-16, H1-1, H1-9 and H1-4, were shown to be abundant in the
 5 subtracted prostate-specific cDNA library. The determined 3' and 5' cDNA sequences for F1-12 are provided in SEQ ID NO: 2 and 3, respectively, with determined 3' cDNA sequences for F1-13, F1-16, H1-1, H1-9 and H1-4 being provided in SEQ ID NO: 1 and 4-7, respectively.

The cDNA sequences for the isolated clones were compared to known
 10 sequences in the gene bank using the EMBL and GenBank databases (release 96). Four of the prostate tumor cDNA clones, F1-13, F1-16, H1-1, and H1-4, were determined to encode the following previously identified proteins: prostate specific antigen (PSA), human glandular kallikrein, human tumor expression enhanced gene, and mitochondria cytochrome C oxidase subunit II. H1-9 was found to be identical to a previously identified
 15 human autonomously replicating sequence. No significant homologies to the cDNA sequence for F1-12 were found.

Subsequent studies led to the isolation of a full-length cDNA sequence for F1-12 (also referred to as P504S). This sequence is provided in SEQ ID NO: 107, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 108. cDNA
 20 splice variants of P504S are provided in SEQ ID NO: 600-605.

To clone less abundant prostate tumor specific genes, cDNA library subtraction was performed by subtracting the prostate tumor cDNA library described above with the normal pancreas cDNA library and with the three most abundant genes in the previously subtracted prostate tumor specific cDNA library: human glandular kallikrein,
 25 prostate specific antigen (PSA), and mitochondria cytochrome C oxidase subunit II. Specifically, 1 µg each of human glandular kallikrein, PSA and mitochondria cytochrome C oxidase subunit II cDNAs in pCDNA3.1 were added to the driver DNA and subtraction was performed as described above to provide a second subtracted cDNA library hereinafter referred to as the "subtracted prostate tumor specific cDNA library with spike".

Twenty-two cDNA clones were isolated from the subtracted prostate tumor specific cDNA library with spike. The determined 3' and 5' cDNA sequences for the clones referred to as J1-17, L1-12, N1-1862, J1-13, J1-19, J1-25, J1-24, K1-58, K1-63, L1-4 and L1-14 are provided in SEQ ID NOS: 8-9, 10-11, 12-13, 14-15, 16-17, 18-19, 20-21, 22-23, 24-25, 26-27 and 28-29, respectively. The determined 3' cDNA sequences for the clones referred to as J1-12, J1-16, J1-21, K1-48, K1-55, L1-2, L1-6, N1-1858, N1-1860, N1-1861, N1-1864 are provided in SEQ ID NOS: 30-40, respectively. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to three of the five most abundant DNA species, (J1-17, L1-12 and N1-1862; SEQ ID NOS: 8-9, 10-11 and 12-13, respectively). Of the remaining two most abundant species, one (J1-12; SEQ ID NO:30) was found to be identical to the previously identified human pulmonary surfactant-associated protein, and the other (K1-48; SEQ ID NO:33) was determined to have some homology to *R. norvegicus* mRNA for 2-arylpropionyl-CoA epimerase. Of the 17 less abundant cDNA clones isolated from the subtracted prostate tumor specific cDNA library with spike, four (J1-16, K1-55, L1-6 and N1-1864; SEQ ID NOS:31, 34, 36 and 40, respectively) were found to be identical to previously identified sequences, two (J1-21 and N1-1860; SEQ ID NOS: 32 and 38, respectively) were found to show some homology to non-human sequences, and two (L1-2 and N1-1861; SEQ ID NOS: 35 and 39, respectively) were found to show some homology to known human sequences. No significant homologies were found to the polypeptides J1-13, J1-19, J1-24, J1-25, K1-58, K1-63, L1-4, L1-14 (SEQ ID NOS: 14-15, 16-17, 20-21, 18-19, 22-23, 24-25, 26-27, 28-29, respectively).

Subsequent studies led to the isolation of full length cDNA sequences for J1-17, L1-12 and N1-1862 (SEQ ID NOS: 109-111, respectively). The corresponding predicted amino acid sequences are provided in SEQ ID NOS: 112-114. L1-12 is also referred to as P501S. A cDNA splice variant of P501S is provided in SEQ ID NO: 606.

In a further experiment, four additional clones were identified by subtracting a prostate tumor cDNA library with normal prostate cDNA prepared from a pool of three normal prostate poly A⁺ RNA (referred to as "prostate subtraction 2"). The determined

cDNA sequences for these clones, hereinafter referred to as U1-3064, U1-3065, V1-3692 and 1A-3905, are provided in SEQ ID NO: 69-72, respectively. Comparison of the determined sequences with those in the gene bank revealed no significant homologies to U1-3065.

5 A second subtraction with spike (referred to as “prostate subtraction spike 2”) was performed by subtracting a prostate tumor specific cDNA library with spike with normal pancreas cDNA library and further spiked with PSA, J1-17, pulmonary surfactant-associated protein, mitochondrial DNA, cytochrome c oxidase subunit II, N1-1862, autonomously replicating sequence, L1-12 and tumor expression enhanced gene. Four
10 additional clones, hereinafter referred to as V1-3686, R1-2330, 1B-3976 and V1-3679, were isolated. The determined cDNA sequences for these clones are provided in SEQ ID NO:73-76, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to V1-3686 and R1-2330.

 Further analysis of the three prostate subtractions described above (prostate
15 subtraction 2, subtracted prostate tumor specific cDNA library with spike, and prostate subtraction spike 2) resulted in the identification of sixteen additional clones, referred to as 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1G-4734, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4810, 1I-4811, 1J-4876, 1K-4884 and 1K-4896. The determined cDNA sequences for these clones are provided in SEQ ID NOS: 77-92, respectively. Comparison
20 of these sequences with those in the gene bank as described above, revealed no significant homologies to 1G-4741, 1G-4734, 1I-4807, 1J-4876 and 1K-4896 (SEQ ID NOS: 79, 81, 87, 90 and 92, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4807, 1J-4876, 1K-4884 and 1K-
25 4896, provided in SEQ ID NOS: 179-188 and 191-193, respectively, and to the determination of additional partial cDNA sequences for 1I-4810 and 1I-4811, provided in SEQ ID NOS: 189 and 190, respectively.

 Additional studies with prostate subtraction spike 2 resulted in the isolation of three more clones. Their sequences were determined as described above and compared

to the most recent GenBank. All three clones were found to have homology to known genes, which are Cysteine-rich protein, KIAA0242, and KIAA0280 (SEQ ID NO: 317, 319, and 320, respectively). Further analysis of these clones by Synteni microarray (Synteni, Palo Alto, CA) demonstrated that all three clones were over-expressed in most prostate tumors and prostate BPH, as well as in the majority of normal prostate tissues tested, but low expression in all other normal tissues.

An additional subtraction was performed by subtracting a normal prostate cDNA library with normal pancreas cDNA (referred to as "prostate subtraction 3"). This led to the identification of six additional clones referred to as 1G-4761, 1G-4762, 1H-4766, 1H-4770, 1H-4771 and 1H-4772 (SEQ ID NOS: 93-98). Comparison of these sequences with those in the gene bank revealed no significant homologies to 1G-4761 and 1H-4771 (SEQ ID NOS: 93 and 97, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4761, 1G-4762, 1H-4766 and 1H-4772 provided in SEQ ID NOS: 194-196 and 199, respectively, and to the determination of additional partial cDNA sequences for 1H-4770 and 1H-4771, provided in SEQ ID NOS: 197 and 198, respectively.

Subtraction of a prostate tumor cDNA library, prepared from a pool of polyA+ RNA from three prostate cancer patients, with a normal pancreas cDNA library (prostate subtraction 4) led to the identification of eight clones, referred to as 1D-4297, 1D-4309, 1D.1-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280 (SEQ ID NOS: 99-107). These sequences were compared to those in the gene bank as described above. No significant homologies were found to 1D-4283 and 1D-4304 (SEQ ID NOS: 103 and 104, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1D-4309, 1D.1-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280, provided in SEQ ID NOS: 200-206, respectively.

cDNA clones isolated in prostate subtraction 1 and prostate subtraction 2, described above, were colony PCR amplified and their mRNA expression levels in prostate tumor, normal prostate and in various other normal tissues were determined using microarray technology (Synteni, Palo Alto, CA). Briefly, the PCR amplification products

were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured.

- 5 This intensity correlates with the hybridization intensity. Two clones (referred to as P509S and P510S) were found to be over-expressed in prostate tumor and normal prostate and expressed at low levels in all other normal tissues tested (liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon). The determined cDNA sequences for
- 10 P509S and P510S are provided in SEQ ID NO: 223 and 224, respectively. Comparison of these sequences with those in the gene bank as described above, revealed some homology to previously identified ESTs.

- Additional, studies led to the isolation of the full-length cDNA sequence for P509S. This sequence is provided in SEQ ID NO: 332, with the corresponding predicted
- 15 amino acid sequence being provided in SEQ ID NO: 339. Two variant full-length cDNA sequences for P510S are provided in SEQ ID NO: 535 and 536, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 537 and 538, respectively. Additional splice variants of P510S are provided in SEQ ID NO: 598 and 599.

- The determined cDNA sequences for additional prostate-specific clones
- 20 isolated during characterization of prostate specific cDNA libraries are provided in SEQ ID NO: 618-689, 691-697 and 709-772. Comparison of these sequences with those in the public databases revealed no significant homologies to any of these sequences.

EXAMPLE 2

25 DETERMINATION OF TISSUE SPECIFICITY OF PROSTATE-SPECIFIC POLYPEPTIDES

Using gene specific primers, mRNA expression levels for the representative prostate-specific polypeptides F1-16, H1-1, J1-17 (also referred to as P502S), L1-12 (also

referred to as P501S), F1-12 (also referred to as P504S) and N1-1862 (also referred to as P503S) were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 1-2 μ g of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR, β -actin was used as an internal control for each of the tissues examined. First, serial dilutions of the first strand cDNAs were prepared and RT-PCR assays were performed using β -actin specific primers. A dilution was then chosen that enabled the linear range amplification of the β -actin template and which was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β -actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding reverse transcriptase.

mRNA Expression levels were examined in four different types of tumor tissue (prostate tumor from 2 patients, breast tumor from 3 patients, colon tumor, lung tumor), and sixteen different normal tissues, including prostate, colon, kidney, liver, lung, ovary, pancreas, skeletal muscle, skin, stomach, testes, bone marrow and brain. F1-16 was found to be expressed at high levels in prostate tumor tissue, colon tumor and normal prostate, and at lower levels in normal liver, skin and testes, with expression being undetectable in the other tissues examined. H1-1 was found to be expressed at high levels in prostate tumor, lung tumor, breast tumor, normal prostate, normal colon and normal brain, at much lower levels in normal lung, pancreas, skeletal muscle, skin, small intestine, bone marrow, and was not detected in the other tissues tested. J1-17 (P502S) and L1-12 (P501S) appear to be specifically over-expressed in prostate, with both genes being expressed at high levels in prostate tumor and normal prostate but at low to undetectable levels in all the other tissues examined. N1-1862 (P503S) was found to be over-expressed in 60% of prostate tumors and detectable in normal colon and kidney. The RT-PCR results

thus indicate that F1-16, H1-1, J1-17 (P502S), N1-1862 (P503S) and L1-12 (P501S) are either prostate specific or are expressed at significantly elevated levels in prostate.

Further RT-PCR studies showed that F1-12 (P504S) is over-expressed in 60% of prostate tumors, detectable in normal kidney but not detectable in all other tissues tested. Similarly, R1-2330 was shown to be over-expressed in 40% of prostate tumors, detectable in normal kidney and liver, but not detectable in all other tissues tested. U1-3064 was found to be over-expressed in 60% of prostate tumors, and also expressed in breast and colon tumors, but was not detectable in normal tissues.

RT-PCR characterization of R1-2330, U1-3064 and 1D-4279 showed that these three antigens are over-expressed in prostate and/or prostate tumors.

Northern analysis with four prostate tumors, two normal prostate samples, two BPH prostates, and normal colon, kidney, liver, lung, pancreas, skeletal muscle, brain, stomach, testes, small intestine and bone marrow, showed that L1-12 (P501S) is over-expressed in prostate tumors and normal prostate, while being undetectable in other normal tissues tested. J1-17 (P502S) was detected in two prostate tumors and not in the other tissues tested. N1-1862 (P503S) was found to be over-expressed in three prostate tumors and to be expressed in normal prostate, colon and kidney, but not in other tissues tested. F1-12 (P504S) was found to be highly expressed in two prostate tumors and to be undetectable in all other tissues tested.

The microarray technology described above was used to determine the expression levels of representative antigens described herein in prostate tumor, breast tumor and the following normal tissues: prostate, liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon. L1-12 (P501S) was found to be over-expressed in normal prostate and prostate tumor, with some expression being detected in normal skeletal muscle. Both J1-12 and F1-12 (P504S) were found to be over-expressed in prostate tumor, with expression being lower or undetectable in all other tissues tested. N1-1862 (P503S) was found to be expressed at high levels in prostate tumor and normal prostate, and at low levels in normal large intestine and normal colon, with expression being undetectable in all

other tissues tested. R1-2330 was found to be over-expressed in prostate tumor and normal prostate, and to be expressed at lower levels in all other tissues tested. 1D-4279 was found to be over-expressed in prostate tumor and normal prostate, expressed at lower levels in normal spinal cord, and to be undetectable in all other tissues tested.

5 Further microarray analysis to specifically address the extent to which P501S (SEQ ID NO: 110) was expressed in breast tumor revealed moderate over-expression not only in breast tumor, but also in metastatic breast tumor (2/31), with negligible to low expression in normal tissues. This data suggests that P501S may be over-expressed in various breast tumors as well as in prostate tumors.

10 The expression levels of 32 ESTs (expressed sequence tags) described by Vasmatzis *et al.* (*Proc. Natl. Acad. Sci. USA* 95:300-304, 1998) in a variety of tumor and normal tissues were examined by microarray technology as described above. Two of these clones (referred to as P1000C and P1001C) were found to be over-expressed in prostate tumor and normal prostate, and expressed at low to undetectable levels in all other tissues
15 tested (normal aorta, thymus, resting and activated PBMC, epithelial cells, spinal cord, adrenal gland, fetal tissues, skin, salivary gland, large intestine, bone marrow, liver, lung, dendritic cells, stomach, lymph nodes, brain, heart, small intestine, skeletal muscle, colon and kidney. The determined cDNA sequences for P1000C and P1001C are provided in SEQ ID NO: 384 and 472, respectively. The sequence of P1001C was found to show some
20 homology to the previously isolated Human mRNA for JM27 protein. No significant homologies were found to the sequence of P1000C.

The expression of the polypeptide encoded by the full length cDNA sequence for F1-12 (also referred to as P504S; SEQ ID NO: 108) was investigated by immunohistochemical analysis. Rabbit-anti-P504S polyclonal antibodies were generated
25 against the full length P504S protein by standard techniques. Subsequent isolation and characterization of the polyclonal antibodies were also performed by techniques well known in the art. Immunohistochemical analysis showed that the P504S polypeptide was expressed in 100% of prostate carcinoma samples tested (n=5).

The rabbit-anti-P504S polyclonal antibody did not appear to label benign prostate cells with the same cytoplasmic granular staining, but rather with light nuclear staining. Analysis of normal tissues revealed that the encoded polypeptide was found to be expressed in some, but not all normal human tissues. Positive cytoplasmic staining with
 5 rabbit-anti-P504S polyclonal antibody was found in normal human kidney, liver, brain, colon and lung-associated macrophages, whereas heart and bone marrow were negative.

This data indicates that the P504S polypeptide is present in prostate cancer tissues, and that there are qualitative and quantitative differences in the staining between benign prostatic hyperplasia tissues and prostate cancer tissues, suggesting that this
 10 polypeptide may be detected selectively in prostate tumors and therefore be useful in the diagnosis of prostate cancer.

EXAMPLE 3

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC

15 POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA subtraction library, containing cDNA from normal prostate subtracted with ten other normal tissue cDNAs (brain, heart, kidney, liver, lung, ovary, placenta, skeletal muscle, spleen and thymus) and then submitted to a first round of PCR
 20 amplification, was purchased from Clontech. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the vector pT7 Blue T-vector (Novagen, Madison, WI) and transformed into XL-1 Blue MRF' *E. coli* (Stratagene). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division
 25 Automated Sequencer Model 373A.

Fifty-nine positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the gene bank, as described above, revealed no significant homologies to 25 of these clones, hereinafter referred to as P5, P8, P9, P18, P20, P30, P34, P36, P38, P39, P42, P49, P50, P53, P55, P60, P64, P65, P73, P75, P76, P79 and

P84. The determined cDNA sequences for these clones are provided in SEQ ID NO: 41-45, 47-52 and 54-65, respectively. P29, P47, P68, P80 and P82 (SEQ ID NO: 46, 53 and 66-68, respectively) were found to show some degree of homology to previously identified DNA sequences. To the best of the inventors' knowledge, none of these sequences have
5 been previously shown to be present in prostate.

Further studies employing the sequence of SEQ ID NO: 67 as a probe in standard full-length cloning methods, resulted in the isolation of three cDNA sequences which appear to be splice variants of P80 (also known as P704P). These sequences are provided in SEQ ID NO: 699-701.

10 Further studies using the PCR-based methodology described above resulted in the isolation of more than 180 additional clones, of which 23 clones were found to show no significant homologies to known sequences. The determined cDNA sequences for these clones are provided in SEQ ID NO: 115-123, 127, 131, 137, 145, 147-151, 153, 156-158 and 160. Twenty-three clones (SEQ ID NO: 124-126, 128-130, 132-136, 138-144, 146,
15 152, 154, 155 and 159) were found to show some homology to previously identified ESTs. An additional ten clones (SEQ ID NO: 161-170) were found to have some degree of homology to known genes. Larger cDNA clones containing the P20 sequence represent splice variants of a gene referred to as P703P. The determined DNA sequence for the variants referred to as DE1, DE13 and DE14 are provided in SEQ ID NOS: 171, 175 and
20 177, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 172, 176 and 178, respectively. The determined cDNA sequence for an extended spliced form of P703 is provided in SEQ ID NO: 225. The DNA sequences for the splice variants referred to as DE2 and DE6 are provided in SEQ ID NOS: 173 and 174, respectively.

25 mRNA Expression levels for representative clones in tumor tissues (prostate (n=5), breast (n=2), colon and lung) normal tissues (prostate (n=5), colon, kidney, liver, lung (n=2), ovary (n=2), skeletal muscle, skin, stomach, small intestine and brain), and activated and non-activated PBMC was determined by RT-PCR as described above. Expression was examined in one sample of each tissue type unless otherwise indicated.

P9 was found to be highly expressed in normal prostate and prostate tumor compared to all normal tissues tested except for normal colon which showed comparable expression. P20, a portion of the P703P gene, was found to be highly expressed in normal prostate and prostate tumor, compared to all twelve normal tissues tested. A modest increase in expression of P20 in breast tumor (n=2), colon tumor and lung tumor was seen compared to all normal tissues except lung (1 of 2). Increased expression of P18 was found in normal prostate, prostate tumor and breast tumor compared to other normal tissues except lung and stomach. A modest increase in expression of P5 was observed in normal prostate compared to most other normal tissues. However, some elevated expression was seen in normal lung and PBMC. Elevated expression of P5 was also observed in prostate tumors (2 of 5), breast tumor and one lung tumor sample. For P30, similar expression levels were seen in normal prostate and prostate tumor, compared to six of twelve other normal tissues tested. Increased expression was seen in breast tumors, one lung tumor sample and one colon tumor sample, and also in normal PBMC. P29 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to the majority of normal tissues. However, substantial expression of P29 was observed in normal colon and normal lung (2 of 2). P80 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to all other normal tissues tested, with increased expression also being seen in colon tumor.

Further studies resulted in the isolation of twelve additional clones, hereinafter referred to as 10-d8, 10-h10, 11-c8, 7-g6, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3, 8-h11, 9-f12 and 9-f3. The determined DNA sequences for 10-d8, 10-h10, 11-c8, 8-d4, 8-d9, 8-h11, 9-f12 and 9-f3 are provided in SEQ ID NO: 207, 208, 209, 216, 217, 220, 221 and 222, respectively. The determined forward and reverse DNA sequences for 7-g6, 8-b5, 8-b6 and 8-g3 are provided in SEQ ID NO: 210 and 211; 212 and 213; 214 and 215; and 218 and 219, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to the sequence of 9-f3. The clones 10-d8, 11-c8 and 8-h11 were found to show some homology to previously isolated ESTs, while 10-h10, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3 and 9-f12 were found to show some homology to previously identified genes.

Further characterization of 7-G6 and 8-G3 showed identity to the known genes PAP and PSA, respectively.

mRNA expression levels for these clones were determined using the micro-array technology described above. The clones 7-G6, 8-G3, 8-B5, 8-B6, 8-D4, 8-D9, 9-F3, 9-F12, 9-H3, 10-A2, 10-A4, 11-C9 and 11-F2 were found to be over-expressed in prostate tumor and normal prostate, with expression in other tissues tested being low or undetectable. Increased expression of 8-F11 was seen in prostate tumor and normal prostate, bladder, skeletal muscle and colon. Increased expression of 10-H10 was seen in prostate tumor and normal prostate, bladder, lung, colon, brain and large intestine. Increased expression of 9-B1 was seen in prostate tumor, breast tumor, and normal prostate, salivary gland, large intestine and skin, with increased expression of 11-C8 being seen in prostate tumor, and normal prostate and large intestine.

An additional cDNA fragment derived from the PCR-based normal prostate subtraction, described above, was found to be prostate specific by both micro-array technology and RT-PCR. The determined cDNA sequence of this clone (referred to as 9-A11) is provided in SEQ ID NO: 226. Comparison of this sequence with those in the public databases revealed 99% identity to the known gene HOXB13.

Further studies led to the isolation of the clones 8-C6 and 8-H7. The determined cDNA sequences for these clones are provided in SEQ ID NO: 227 and 228, respectively. These sequences were found to show some homology to previously isolated ESTs.

PCR and hybridization-based methodologies were employed to obtain longer cDNA sequences for clone P20 (also referred to as P703P), yielding three additional cDNA fragments that progressively extend the 5' end of the gene. These fragments, referred to as P703PDE5, P703P6.26, and P703PX-23 (SEQ ID NO: 326, 328 and 330, with the predicted corresponding amino acid sequences being provided in SEQ ID NO: 327, 329 and 331, respectively) contain additional 5' sequence. P703PDE5 was recovered by screening of a cDNA library (#141-26) with a portion of P703P as a probe. P703P6.26 was recovered from a mixture of three prostate tumor cDNAs and P703PX_23 was

recovered from cDNA library (#438-48). Together, the additional sequences include all of the putative mature serine protease along with part of the putative signal sequence. The full-length cDNA sequence for P703P is provided in SEQ ID NO: 524, with the corresponding amino acid sequence being provided in SEQ ID NO: 525.

5 Using computer algorithms, the following regions of P703P were predicted to represent potential HLA A2-binding CTL epitopes: amino acids 164-172 of SEQ ID NO: 525 (SEQ ID NO: 866); amino acids 160-168 of SEQ ID NO: 525 (SEQ ID NO: 867); amino acids 239-247 of SEQ ID NO: 525 (SEQ ID NO: 868); amino acids 118-126 of SEQ ID NO: 525 (SEQ ID NO: 869); amino acids 112-120 of SEQ ID NO: 525 (SEQ ID NO: 870); amino acids 155-164 of SEQ ID NO: 525 (SEQ ID NO: 871); amino acids 117-126 of SEQ ID NO: 525 (SEQ ID NO: 872); amino acids 164-173 of SEQ ID NO: 525 (SEQ ID NO: 873); amino acids 154-163 of SEQ ID NO: 525 (SEQ ID NO: 874); amino acids 163-172 of SEQ ID NO: 525 (SEQ ID NO: 875); amino acids 58-66 of SEQ ID NO: 525 (SEQ ID NO: 876); and amino acids 59-67 of SEQ ID NO: 525 (SEQ ID NO: 877).

15 P703P was found to show some homology to previously identified proteases, such as thrombin. The thrombin receptor has been shown to be preferentially expressed in highly metastatic breast carcinoma cells and breast carcinoma biopsy samples. Introduction of thrombin receptor antisense cDNA has been shown to inhibit the invasion of metastatic breast carcinoma cells in culture. Antibodies against thrombin receptor
20 inhibit thrombin receptor activation and thrombin-induced platelet activation. Furthermore, peptides that resemble the receptor's tethered ligand domain inhibit platelet aggregation by thrombin. P703P may play a role in prostate cancer through a protease-activated receptor on the cancer cell or on stromal cells. The potential trypsin-like protease activity of P703P may either activate a protease-activated receptor on the cancer cell membrane to promote
25 tumorigenesis or activate a protease-activated receptor on the adjacent cells (such as stromal cells) to secrete growth factors and/or proteases (such as matrix metalloproteinases) that could promote tumor angiogenesis, invasion and metastasis. P703P may thus promote tumor progression and/or metastasis through the activation of protease-activated receptor.

Polypeptides and antibodies that block the P703P-receptor interaction may therefore be usefully employed in the treatment of prostate cancer.

To determine whether P703P expression increases with increased severity of Gleason grade, an indicator of tumor stage, quantitative PCR analysis was performed on prostate tumor samples with a range of Gleason scores from 5 to > 8. The mean level of P703P expression increased with increasing Gleason score, indicating that P703P expression may correlate with increased disease severity.

Further studies using a PCR-based subtraction library of a prostate tumor pool subtracted against a pool of normal tissues (referred to as JP: PCR subtraction) resulted in the isolation of thirteen additional clones, seven of which did not share any significant homology to known GenBank sequences. The determined cDNA sequences for these seven clones (P711P, P712P, novel 23, P774P, P775P, P710P and P768P) are provided in SEQ ID NO: 307-311, 313 and 315, respectively. The remaining six clones (SEQ ID NO: 316 and 321-325) were shown to share some homology to known genes. By microarray analysis, all thirteen clones showed three or more fold over-expression in prostate tissues, including prostate tumors, BPH and normal prostate as compared to normal non-prostate tissues. Clones P711P, P712P, novel 23 and P768P showed over-expression in most prostate tumors and BPH tissues tested (n=29), and in the majority of normal prostate tissues (n=4), but background to low expression levels in all normal tissues. Clones P774P, P775P and P710P showed comparatively lower expression and expression in fewer prostate tumors and BPH samples, with negative to low expression in normal prostate.

Further studies led to the isolation of an extended cDNA sequence for P712P (SEQ ID NO: 552). The amino acid sequences encoded by 16 predicted open reading frames present within the sequence of SEQ ID NO: 552 are provided in SEQ ID NO: 553-568.

The full-length cDNA for P711P was obtained by employing the partial sequence of SEQ ID NO: 307 to screen a prostate cDNA library. Specifically, a directionally cloned prostate cDNA library was prepared using standard techniques. One

million colonies of this library were plated onto LB/Amp plates. Nylon membrane filters were used to lift these colonies, and the cDNAs which were picked up by these filters were denatured and cross-linked to the filters by UV light. The P711P cDNA fragment of SEQ ID NO: 307 was radio-labeled and used to hybridize with these filters. Positive clones
5 were selected, and cDNAs were prepared and sequenced using an automatic Perkin Elmer/Applied Biosystems sequencer. The determined full-length sequence of P711P is provided in SEQ ID NO: 382, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 383.

Using PCR and hybridization-based methodologies, additional cDNA
10 sequence information was derived for two clones described above, 11-C9 and 9-F3, herein after referred to as P707P and P714P, respectively (SEQ ID NO: 333 and 334). After comparison with the most recent GenBank, P707P was found to be a splice variant of the known gene HoxB13. In contrast, no significant homologies to P714P were found. Further studies employing the sequence of SEQ ID NO: 334 as a probe in standard full-length
15 cloning methods, resulted in an extended cDNA sequence for P714P. This sequence is provided in SEQ ID NO: 698. This sequence was found to show some homology to the gene that encodes human ribosomal L23A protein.

Clones 8-B3, P89, P98, P130 and P201 (as disclosed in U.S. Patent Application No. 09/020,956, filed February 9, 1998) were found to be contained within one
20 contiguous sequence, referred to as P705P (SEQ ID NO: 335, with the predicted amino acid sequence provided in SEQ ID NO: 336), which was determined to be a splice variant of the known gene NKX 3.1.

Further studies on P775P resulted in the isolation of four additional sequences (SEQ ID NO: 473-476) which are all splice variants of the P775P gene. The
25 sequence of SEQ ID NO: 474 was found to contain two open reading frames (ORFs). The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 477 and 478. The cDNA sequence of SEQ ID NO: 475 was found to contain an ORF which encodes the amino acid sequence of SEQ ID NO: 479. The cDNA sequence of SEQ ID NO: 473 was found to contain four ORFs. The predicted amino acid sequences encoded by

these ORFs are provided in SEQ ID NO: 480-483. Additional splice variants of P775P are provided in SEQ ID NO: 593-597.

Subsequent studies led to the identification of a genomic region on chromosome 22q11.2, known as the Cat Eye Syndrome region, that contains the five prostate genes P704P, P712P, P774P, P775P and B305D. The relative location of each of these five genes within the genomic region is shown in Fig. 10. This region may therefore be associated with malignant tumors, and other potential tumor genes may be contained within this region. These studies also led to the identification of a potential open reading frame (ORF) for P775P (provided in SEQ ID NO: 533), which encodes the amino acid sequence of SEQ ID NO: 534.

Comparison of the clone of SEQ ID NO: 325 (referred to as P558S) with sequences in the GenBank and GeneSeq DNA databases showed that P558S is identical to the prostate-specific transglutaminase gene, which is known to have two forms. The full-length sequences for the two forms are provided in SEQ ID NO: 773 and 774, with the corresponding amino acid sequences being provided in SEQ ID NO: 775 and 776, respectively. The cDNA sequence of SEQ ID NO: 774 has a 15 pair base insert, resulting in a 5 amino acid insert in the corresponding amino acid sequence (SEQ ID NO: 776). This insert is not present in the sequence of SEQ ID NO: 773.

EXAMPLE 4

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems 430A peptide synthesizer using Fmoc chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the

peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following
 5 lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

EXAMPLE 5

FURTHER ISOLATION AND CHARACTERIZATION OF

10 PROSTATE-SPECIFIC POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA library generated from prostate primary tumor mRNA as described above was subtracted with cDNA from normal prostate. The subtraction was performed using a PCR-based protocol (Clontech), which was modified to generate larger fragments.
 15 Within this protocol, tester and driver double stranded cDNA were separately digested with five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the Clontech protocol. This modification did not affect the subtraction efficiency. Two tester
 20 populations were then created with different adapters, and the driver library remained without adapters.

The tester and driver libraries were then hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester
 25 cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized

to tester cDNA with the second adaptor. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using
5 adaptor-specific primers. Only population (e), which contained tester cDNA that did not hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich differentially expressed sequences.

This PCR-based subtraction technique normalizes differentially expressed
10 cDNAs so that rare transcripts that are overexpressed in prostate tumor tissue may be recoverable. Such transcripts would be difficult to recover by traditional subtraction methods.

In addition to genes known to be overexpressed in prostate tumor, seventy-seven further clones were identified. Sequences of these partial cDNAs are provided in
15 SEQ ID NO: 29 to 305. Most of these clones had no significant homology to database sequences. Exceptions were JTPN23 (SEQ ID NO: 231; similarity to pig valosin-containing protein), JTPN30 (SEQ ID NO: 234; similarity to rat mRNA for proteasome subunit), JTPN45 (SEQ ID NO: 243; similarity to rat *norvegicus* cytosolic NADP-dependent isocitrate dehydrogenase), JTPN46 (SEQ ID NO: 244; similarity to human
20 subclone H8 4 d4 DNA sequence), JP1D6 (SEQ ID NO: 265; similarity to *G. gallus* dynein light chain-A), JP8D6 (SEQ ID NO: 288; similarity to human BAC clone RG016J04), JP8F5 (SEQ ID NO: 289; similarity to human subclone H8 3 b5 DNA sequence), and JP8E9 (SEQ ID NO: 299; similarity to human Alu sequence).

Additional studies using the PCR-based subtraction library consisting of a
25 prostate tumor pool subtracted against a normal prostate pool (referred to as PT-PN PCR subtraction) yielded three additional clones. Comparison of the cDNA sequences of these clones with the most recent release of GenBank revealed no significant homologies to the two clones referred to as P715P and P767P (SEQ ID NO: 312 and 314). The remaining clone was found to show some homology to the known gene KIAA0056 (SEQ ID NO:

318). Using microarray analysis to measure mRNA expression levels in various tissues, all three clones were found to be over-expressed in prostate tumors and BPH tissues. Specifically, clone P715P was over-expressed in most prostate tumors and BPH tissues by a factor of three or greater, with elevated expression seen in the majority of normal prostate samples and in fetal tissue, but negative to low expression in all other normal tissues. Clone P767P was over-expressed in several prostate tumors and BPH tissues, with moderate expression levels in half of the normal prostate samples, and background to low expression in all other normal tissues tested.

Further analysis, by microarray as described above, of the PT-PN PCR subtraction library and of a DNA subtraction library containing cDNA from prostate tumor subtracted with a pool of normal tissue cDNAs, led to the isolation of 27 additional clones (SEQ ID NO: 340-365 and 381) which were determined to be over-expressed in prostate tumor. The clones of SEQ ID NO: 341, 342, 345, 347, 348, 349, 351, 355-359, 361, 362 and 364 were also found to be expressed in normal prostate. Expression of all 26 clones in a variety of normal tissues was found to be low or undetectable, with the exception of P544S (SEQ ID NO: 356) which was found to be expressed in small intestine. Of the 26 clones, 11 (SEQ ID NO: 340-349 and 362) were found to show some homology to previously identified sequences. No significant homologies were found to the clones of SEQ ID NO: 350, 351, 353-361, and 363-365.

Comparison of the sequence of SEQ ID NO: 362 with sequences in the GenBank and GeneSeq DNA databases showed that this clone (referred to as P788P) is identical to GeneSeq Accession No. X27262, which encodes a protein found in the GeneSeq protein Accession No. Y00931. The full length cDNA sequence of P788P is shown in Figure 12A (SEQ ID NO: 777), with the corresponding predicted amino acid being shown in Figure 12B (SEQ ID NO: 778). Subsequently, a full-length cDNA sequence for P788P that contains polymorphisms not found in the sequence of SEQ ID NO: 779, was cloned multiple times by PCR amplification from cDNA prepared from several RNA templates from three individuals. This determined cDNA sequence of this polymorphic variant of P788P is provided in SEQ ID NO: 779, with the corresponding

amino acid sequence being provided in SEQ ID NO: 780. The sequence of SEQ ID NO: 780 differs from that of SEQ ID NO: 778 by six amino acid residues. The P788P protein has 7 potential transmembrane domains at the C-terminal portion and is predicted to be a plasma membrane protein with an extracellular N-terminal region.

5 Further studies on the clone of SEQ ID NO: 352 (referred to as P790P) led to the isolation of the full-length cDNA sequence of SEQ ID NO: 526. The corresponding predicted amino acid is provided in SEQ ID NO: 527. Data from two quantitative PCR experiments indicated that P790P is over-expressed in 11/15 tested prostate tumor samples and is expressed at low levels in spinal cord, with no expression being seen in all other
10 normal samples tested. Data from further PCR experiments and microarray experiments showed over-expression in normal prostate and prostate tumor with little or no expression in other tissues tested. P790P was subsequently found to show significant homology to a previously identified G-protein coupled prostate tissue receptor.

Additional studies on the clone of SEQ ID NO: 354 (referred to as P776P)
15 led to the isolation of an extended cDNA sequence, provided in SEQ ID NO: 569. The determined cDNA sequences of three additional splice variants of P776P are provided in SEQ ID NO: 570-572. The amino acid sequences encoded by two predicted open reading frames (ORFs) contained within SEQ ID NO: 570, one predicted ORF contained within SEQ ID NO: 571, and 11 predicted ORFs contained within SEQ ID NO: 569, are provided
20 in SEQ ID NO: 573-586, respectively.

Comparison of the cDNA sequences for the clones P767P (SEQ ID NO: 314) and P777P (SEQ ID NO: 350) with sequences in the GenBank human EST database showed that the two clones matched many EST sequences in common, suggesting that P767P and P777P may represent the same gene. A DNA consensus sequence derived from
25 a DNA sequence alignment of P767P, P777P and multiple EST clones is provided in SEQ ID NO: 587. The amino acid sequences encoded by three putative ORFs located within SEQ ID NO: 587 are provided in SEQ ID NO: 588-590.

EXAMPLE 6

PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

6.1. This Example illustrates the preparation of a CTL cell line specific for
5 cells expressing the P502S gene.

Mice expressing the transgene for human HLA A2Kb (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with P2S#12 peptide (VLGWVAEL; SEQ ID NO: 306), which is derived from the P502S gene (also referred to herein as J1-17, SEQ ID NO: 8), as described by Theobald et al., *Proc. Natl.*
10 *Acad. Sci. USA* 92:11993-11997, 1995 with the following modifications. Mice were immunized with 100µg of P2S#12 and 120µg of an I-A^b binding peptide derived from hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and using a nylon mesh single cell suspensions prepared. Cells were then resuspended at 6×10^6 cells/ml in complete media (RPMI-1640; Gibco BRL,
15 Gaithersburg, MD) containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL), non-essential amino acids (Gibco BRL), 2×10^{-5} M 2-mercaptoethanol, 50U/ml penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads) P2S#12-pulsed (5mg/ml P2S#12 and 10mg/ml β 2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7µg/ml dextran sulfate and 25µg/ml
20 LPS for 3 days). Six days later, cells (5×10^5 /ml) were restimulated with 2.5×10^6 /ml peptide pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman et al, *Science* 258:815-818, 1992) and 3×10^6 /ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20U/ml IL-2. Cells continued to be restimulated on a weekly basis as described, in preparation for cloning the line.

25 P2S#12 line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1×10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5×10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, clones that were growing were isolated and maintained in culture. Several of these clones demonstrated significantly higher reactivity (lysis)

against human fibroblasts (HLA A2Kb expressing) transduced with P502S than against control fibroblasts. An example is presented in Figure 1.

This data indicates that P2S #12 represents a naturally processed epitope of the P502S protein that is expressed in the context of the human HLA A2Kb molecule.

5

6.2. This Example illustrates the preparation of murine CTL lines and CTL clones specific for cells expressing the P501S gene.

This series of experiments were performed similarly to that described above.

10 Mice were immunized with the P1S#10 peptide (SEQ ID NO: 337), which is derived from the P501S gene (also referred to herein as L1-12, SEQ ID NO: 110). The P1S#10 peptide was derived by analysis of the predicted polypeptide sequence for P501S for potential HLA-A2 binding sequences as defined by published HLA-A2 binding motifs (Parker, KC, *et al*, *J. Immunol.*, 152:163, 1994). P1S#10 peptide was synthesized as described in

15 Example 4, and empirically tested for HLA-A2 binding using a T cell based competition assay. Predicted A2 binding peptides were tested for their ability to compete HLA-A2 specific peptide presentation to an HLA-A2 restricted CTL clone (D150M58), which is specific for the HLA-A2 binding influenza matrix peptide fluM58. D150M58 CTL secretes TNF in response to self-presentation of peptide fluM58. In the competition assay,

20 test peptides at 100-200 µg/ml were added to cultures of D150M58 CTL in order to bind HLA-A2 on the CTL. After thirty minutes, CTL cultured with test peptides, or control peptides, were tested for their antigen dose response to the fluM58 peptide in a standard TNF bioassay. As shown in Figure 3, peptide P1S#10 competes HLA-A2 restricted presentation of fluM58, demonstrating that peptide P1S#10 binds HLA-A2.

25 Mice expressing the transgene for human HLA A2Kb were immunized as described by Theobald et al. (*Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995) with the following modifications. Mice were immunized with 62.5µg of P1S #10 and 120µg of an I-A^b binding peptide derived from Hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and single cell

suspensions prepared using a nylon mesh. Cells were then resuspended at 6×10^6 cells/ml in complete media (as described above) and cultured in the presence of irradiated (3000 rads) P1S#10-pulsed ($2\mu\text{g/ml}$ P1S#10 and 10mg/ml $\beta 2$ -microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of $7\mu\text{g/ml}$ dextran sulfate and $25\mu\text{g/ml}$ LPS for 3 days). Six days later cells ($5 \times 10^5/\text{ml}$) were restimulated with $2.5 \times 10^6/\text{ml}$ peptide-pulsed irradiated (20,000 rads) EL4A2Kb cells, as described above, and $3 \times 10^6/\text{ml}$ A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20 U/ml IL-2. Cells were restimulated on a weekly basis in preparation for cloning. After three rounds of *in vitro* stimulations, one line was generated that recognized P1S#10-pulsed Jurkat A2Kb targets and P501S-transduced Jurkat targets as shown in Figure 4.

A P1S#10-specific CTL line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1×10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5×10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, viable clones were isolated and maintained in culture. As shown in Figure 5, five of these clones demonstrated specific cytolytic reactivity against P501S-transduced Jurkat A2Kb targets. This data indicates that P1S#10 represents a naturally processed epitope of the P501S protein that is expressed in the context of the human HLA-A2.1 molecule.

EXAMPLE 7

PRIMING OF CTL *IN VIVO* USING NAKED DNA IMMUNIZATION

WITH A PROSTATE ANTIGEN

The prostate-specific antigen L1-12, as described above, is also referred to as P501S. HLA A2Kb Tg mice (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with $100 \mu\text{g}$ P501S in the vector VR1012 either intramuscularly or intradermally. The mice were immunized three times, with a two week interval between immunizations. Two weeks after the last immunization, immune spleen cells were cultured with Jurkat A2Kb-P501S transduced stimulator cells. CTL lines were stimulated weekly. After two weeks of *in vitro* stimulation, CTL activity was assessed

against P501S transduced targets. Two out of 8 mice developed strong anti-P501S CTL responses. These results demonstrate that P501S contains at least one naturally processed HLA-A2-restricted CTL epitope.

5

EXAMPLE 8

ABILITY OF HUMAN T CELLS TO RECOGNIZE PROSTATE-SPECIFIC POLYPEPTIDES

This Example illustrates the ability of T cells specific for a prostate tumor polypeptide to recognize human tumor.

10

Human CD8⁺ T cells were primed *in vitro* to the P2S-12 peptide (SEQ ID NO: 306) derived from P502S (also referred to as J1-17) using dendritic cells according to the protocol of Van Tsai et al. (*Critical Reviews in Immunology* 18:65-75, 1998). The resulting CD8⁺ T cell microcultures were tested for their ability to recognize the P2S-12 peptide presented by autologous fibroblasts or fibroblasts which were transduced to express the P502S gene in a γ -interferon ELISPOT assay (*see* Lalvani et al., *J. Exp. Med.* 186:859-865, 1997). Briefly, titrating numbers of T cells were assayed in duplicate on 10⁴ fibroblasts in the presence of 3 μ g/ml human β_2 -microglobulin and 1 μ g/ml P2S-12 peptide or control E75 peptide. In addition, T cells were simultaneously assayed on autologous fibroblasts transduced with the P502S gene or as a control, fibroblasts transduced with HER-2/*neu*. Prior to the assay, the fibroblasts were treated with 10 ng/ml γ -interferon for 48 hours to upregulate class I MHC expression. One of the microcultures (#5) demonstrated strong recognition of both peptide pulsed fibroblasts as well as transduced fibroblasts in a γ -interferon ELISPOT assay. Figure 2A demonstrates that there was a strong increase in the number of γ -interferon spots with increasing numbers of T cells on fibroblasts pulsed with the P2S-12 peptide (solid bars) but not with the control E75 peptide (open bars). This shows the ability of these T cells to specifically recognize the P2S-12 peptide. As shown in Figure 2B, this microculture also demonstrated an increase in the number of γ -interferon spots with increasing numbers of T cells on fibroblasts transduced to express the P502S gene but not the HER-2/*neu* gene. These results provide additional

15

20

25

confirmatory evidence that the P2S-12 peptide is a naturally processed epitope of the P502S protein. Furthermore, this also demonstrates that there exists in the human T cell repertoire, high affinity T cells which are capable of recognizing this epitope. These T cells should also be capable of recognizing human tumors which express the P502S gene.

5

EXAMPLE 9

ELICITATION OF PROSTATE ANTIGEN-SPECIFIC CTL RESPONSES

IN HUMAN BLOOD

10 This Example illustrates the ability of a prostate-specific antigen to elicit a CTL response in blood of normal humans.

Autologous dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal donors by growth for five days in RPMI medium containing 10% human serum, 50 ng/ml GMCSF and 30 ng/ml IL-4. Following culture, DC were
 15 infected overnight with recombinant P501S-expressing vaccinia virus at an M.O.I. of 5 and matured for 8 hours by the addition of 2 micrograms/ml CD40 ligand. Virus was inactivated by UV irradiation, CD8⁺ cells were isolated by positive selection using magnetic beads, and priming cultures were initiated in 24-well plates. Following five stimulation cycles using autologous fibroblasts retrovirally transduced to express P501S
 20 and CD80, CD8⁺ lines were identified that specifically produced interferon-gamma when stimulated with autologous P501S-transduced fibroblasts. The P501S-specific activity of cell line 3A-1 could be maintained following additional stimulation cycles on autologous B-LCL transduced with P501S. Line 3A-1 was shown to specifically recognize autologous B-LCL transduced to express P501S, but not EGFP-transduced autologous B-LCL, as
 25 measured by cytotoxicity assays (⁵¹Cr release) and interferon-gamma production (Interferon-gamma Elispot; *see above and Lalvani et al., J. Exp. Med.* 186:859-865, 1997). The results of these assays are presented in Figures 6A and 6B.

EXAMPLE 10

IDENTIFICATION OF A NATURALLY PROCESSED CTL EPITOPE CONTAINED WITHIN THE
PROSTATE-SPECIFIC ANTIGEN P703P

5 The 9-mer peptide p5 (SEQ ID NO: 338) was derived from the P703P antigen (also referred to as P20). The p5 peptide is immunogenic in human HLA-A2 donors and is a naturally processed epitope. Antigen specific human CD8+ T cells can be primed following repeated *in vitro* stimulations with monocytes pulsed with p5 peptide. These CTL specifically recognize p5-pulsed and P703P-transduced target cells in both
10 ELISPOT (as described above) and chromium release assays. Additionally, immunization of HLA-A2Kb transgenic mice with p5 leads to the generation of CTL lines which recognize a variety of HLA-A2Kb or HLA-A2 transduced target cells expressing P703P.

 Initial studies demonstrating that p5 is a naturally processed epitope were done using HLA-A2Kb transgenic mice. HLA-A2Kb transgenic mice were immunized
15 subcutaneously in the footpad with 100 µg of p5 peptide together with 140 µg of hepatitis B virus core peptide (a Th peptide) in Freund's incomplete adjuvant. Three weeks post immunization, spleen cells from immunized mice were stimulated *in vitro* with peptide-pulsed LPS blasts. CTL activity was assessed by chromium release assay five days after primary *in vitro* stimulation. Retrovirally transduced cells expressing the control antigen
20 P703P and HLA-A2Kb were used as targets. CTL lines that specifically recognized both p5-pulsed targets as well as P703P-expressing targets were identified.

 Human *in vitro* priming experiments demonstrated that the p5 peptide is immunogenic in humans. Dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by culturing for five days in RPMI medium
25 containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, the DC were pulsed with 1 ug/ml p5 peptide and cultured with CD8+ T cell enriched PBMC. CTL lines were restimulated on a weekly basis with p5-pulsed monocytes. Five to six weeks after initiation of the CTL cultures, CTL recognition of p5-

pulsed target cells was demonstrated. CTL were additionally shown to recognize human cells transduced to express P703P, demonstrating that p5 is a naturally processed epitope.

Studies identifying a further peptide epitope (referred to as peptide 4) derived from the prostate tumor-specific antigen P703P that is capable of being recognized by CD4 T cells on the surface of cells in the context of HLA class II molecules were carried out as follows. The amino acid sequence for peptide 4 is provided in SEQ ID NO: 781, with the corresponding cDNA sequence being provided in SEQ ID NO: 782.

Twenty 15-mer peptides overlapping by 10 amino acids and derived from the carboxy-terminal fragment of P703P were generated using standard procedures. Dendritic cells (DC) were derived from PBMC of a normal female donor using GM-CSF and IL-4 by standard protocols. CD4 T cells were generated from the same donor as the DC using MACS beads and negative selection. DC were pulsed overnight with pools of the 15-mer peptides, with each peptide at a final concentration of 0.25 microgram/ml. Pulsed DC were washed and plated at 1×10^4 cells/well of 96-well V-bottom plates and purified CD4 T cells were added at 1×10^5 /well. Cultures were supplemented with 60 ng/ml IL-6 and 10 ng/ml IL-12 and incubated at 37 °C. Cultures were restimulated as above on a weekly basis using DC generated and pulsed as above as antigen presenting cells, supplemented with 5 ng/ml IL-7 and 10 u/ml IL-2. Following 4 *in vitro* stimulation cycles, 96 lines (each line corresponding to one well) were tested for specific proliferation and cytokine production in response to the stimulating pools with an irrelevant pool of peptides derived from mammaglobin being used as a control.

One line (referred to as 1-F9) was identified from pool #1 that demonstrated specific proliferation (measured by 3H proliferation assays) and cytokine production (measured by interferon-gamma ELISA assays) in response to pool #1 of P703P peptides. This line was further tested for specific recognition of the peptide pool, specific recognition of individual peptides in the pool, and in HLA mismatch analyses to identify the relevant restricting allele. Line 1-F9 was found to specifically proliferate and produce interferon-gamma in response to peptide pool #1, and also to peptide 4 (SEQ ID NO: 781). Peptide 4 corresponds to amino acids 126-140 of SEQ ID NO: 327. Peptide titration experiments

were conducted to assess the sensitivity of line 1-F9 for the specific peptide. The line was found to specifically respond to peptide 4 at concentrations as low as 0.25 ng/ml, indicating that the T cells are very sensitive and therefore likely to have high affinity for the epitope.

To determine the HLA restriction of the P703P response, a panel of antigen presenting cells (APC) was generated that was partially matched with the donor used to generate the T cells. The APC were pulsed with the peptide and used in proliferation and cytokine assays together with line 1-F9. APC matched with the donor at HLA-DRB0701 and HLA-DQB02 alleles were able to present the peptide to the T cells, indicating that the P703P-specific response is restricted to one of these alleles.

Antibody blocking assays were utilized to determine if the restricting allele was HLA-DR0701 or HLA-DQ02. The anti-HLA-DR blocking antibody L243 or an irrelevant isotype matched IgG2a were added to T cells and APC cultures pulsed with the peptide RMPTVLQCVNVS VVS (SEQ ID NO: 781) at 250 ng/ml. Standard interferon-gamma and proliferation assays were performed. Whereas the control antibody had no effect on the ability of the T cells to recognize peptide-pulsed APC, in both assays the anti-HLA-DR antibody completely blocked the ability of the T cells to specifically recognize peptide-pulsed APC.

To determine if the peptide epitope RMPTVLQCVNVS VVS (SEQ ID NO: 781) was naturally processed, the ability of line 1-F9 to recognize APC pulsed with recombinant P703P protein was examined. For these experiments a number of recombinant P703P sources were utilized; *E. coli*-derived P703P, Pichia-derived P703P and baculovirus-derived P703P. Irrelevant protein controls used were *E. coli*-derived L3E (a lung-specific antigen) and baculovirus-derived mammaglobin. In interferon-gamma ELISA assays, line 1-F9 was able to efficiently recognize both *E. coli* forms of P703P as well as Pichia-derived recombinant P703P, while baculovirus-derived P703P was recognized less efficiently. Subsequent Western blot analysis revealed that the *E. coli* and Pichia P703P protein preparations were intact while the baculovirus P703P preparation was approximately 75% degraded. Thus, peptide RMPTVLQCVNVS VVS (SEQ ID NO: 781)

from P703P is a naturally processed peptide epitope derived from P703P and presented to T cells in the context of HLA-DRB-0701

In further studies, twenty-four 15-mer peptides overlapping by 10 amino acids and derived from the N-terminal fragment of P703P (corresponding to amino acids 27-154 of SEQ ID NO: 525) were generated by standard procedures and their ability to be recognized by CD4 cells was determined essentially as described above. DC were pulsed overnight with pools of the peptides with each peptide at a final concentration of 10 microgram/ml. A large number of individual CD4 T cell lines (65/480) demonstrated significant proliferation and cytokine release (IFN-gamma) in response to the P703P peptide pools but not to a control peptide pool. The CD4 T cell lines which demonstrated specific activity were restimulated on the appropriate pool of P703P peptides and reassayed on the individual peptides of each pool as well as a peptide dose titration of the pool of peptides in a IFN-gamma release assay and in a proliferation assay.

Sixteen immunogenic peptides were recognized by the T cells from the entire set of peptide antigens tested. The amino acid sequences of these peptides are provided in SEQ ID NO: 799-814, with the corresponding cDNA sequences being provided in SEQ ID NO: 783-798, respectively. In some cases the peptide reactivity of the T cell line could be mapped to a single peptide, however some could be mapped to more than one peptide in each pool. Those CD4 T cell lines that displayed a representative pattern of recognition from each peptide pool with a reasonable affinity for peptide were chosen for further analysis (I-1A, -6A; II-4C, -5E; III-6E, IV-4B, -3F, -9B, -10F, V-5B, -4D, and -10F). These CD4 T cells lines were restimulated on the appropriate individual peptide and reassayed on autologous DC pulsed with a truncated form of recombinant P703P protein made in *E. coli* (a.a. 96 - 254 of SEQ ID NO: 525), full-length P703P made in the baculovirus expression system, and a fusion between influenza virus NS1 and P703P made in *E. coli*. Of the T cell lines tested, line I-1A recognized specifically the truncated form of P703P (*E. coli*) but no other recombinant form of P703P. This line also recognized the peptide used to elicit the T cells. Line 2-4C recognized the truncated form of P703P (*E. coli*) and the full length form of P703P made in baculovirus, as well as peptide. The

remaining T cell lines tested were either peptide-specific only (II-5E, II-6F, IV-4B, IV-3F, IV-9B, IV-10F, V-5B and V-4D) or were non-responsive to any antigen tested (V-10F). These results demonstrate that the peptide sequence RPLLANDLMLIKLDE (SEQ ID NO: 814; corresponding to a.a. 110-124 of SEQ ID NO: 525) recognized by the T cell line I-1A, and the peptide sequences SVSESDTIRSISIAS (SEQ ID NO: 811; corresponding to a.a. 125-139 of SEQ ID NO: 525) and ISIASQCPTAGNSCL (SEQ ID NO: 810; corresponding to a.a. 135-149 of SEQ ID NO: 525) recognized by the T cell line II-4C may be naturally processed epitopes of the P703P protein.

10

EXAMPLE 11

EXPRESSION OF A BREAST TUMOR-DERIVED ANTIGEN
IN PROSTATE

Isolation of the antigen B305D from breast tumor by differential display is described in US Patent Application No. 08/700,014, filed August 20, 1996. Several different splice forms of this antigen were isolated. The determined cDNA sequences for these splice forms are provided in SEQ ID NO: 366-375, with the predicted amino acid sequences corresponding to the sequences of SEQ ID NO: 292, 298 and 301-303 being provided in SEQ ID NO: 299-306, respectively. In further studies, a splice variant of the cDNA sequence of SEQ ID NO: 366 was isolated which was found to contain an additional guanine residue at position 884 (SEQ ID NO: 530), leading to a frameshift in the open reading frame. The determined DNA sequence of this ORF is provided in SEQ ID NO: 531. This frameshift generates a protein sequence (provided in SEQ ID NO: 532) of 293 amino acids that contains the C-terminal domain common to the other isoforms of B305D but that differs in the N-terminal region.

The expression levels of B305D in a variety of tumor and normal tissues were examined by real time PCR and by Northern analysis. The results indicated that B305D is highly expressed in breast tumor, prostate tumor, normal prostate and normal testes, with expression being low or undetectable in all other tissues examined (colon

tumor, lung tumor, ovary tumor, and normal bone marrow, colon, kidney, liver, lung, ovary, skin, small intestine, stomach). Using real-time PCR on a panel of prostate tumors, expression of B305D in prostate tumors was shown to increase with increasing Gleason grade, demonstrating that expression of B305D increases as prostate cancer progresses.

5

EXAMPLE 12

GENERATION OF HUMAN CTL *IN VITRO* USING WHOLE GENE PRIMING AND STIMULATION TECHNIQUES WITH THE PROSTATE-SPECIFIC ANTIGEN P501S

10 Using *in vitro* whole-gene priming with P501S-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with P501S (also known as L1-12), as determined by interferon- γ ELISPOT analysis as described above. Using a panel of HLA-mismatched B-LCL lines transduced with P501S, these CTL

15 lines were shown to be likely restricted to HLAB class I allele. Specifically, dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC were infected overnight with recombinant P501S vaccinia virus at a multiplicity of infection (M.O.I) of five, and

20 matured overnight by the addition of 3 μ g/ml CD40 ligand. Virus was inactivated by UV irradiation. CD8⁺ T cells were isolated using a magnetic bead system, and priming cultures were initiated using standard culture techniques. Cultures were restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with P501S and CD80. Following four stimulation cycles, CD8⁺ T cell lines were identified that

25 specifically produced interferon- γ when stimulated with P501S and CD80-transduced autologous fibroblasts. A panel of HLA-mismatched B-LCL lines transduced with P501S were generated to define the restriction allele of the response. By measuring interferon- γ in an ELISPOT assay, the P501S specific response was shown to be likely restricted by HLA B alleles. These results demonstrate that a CD8⁺ CTL response to P501S can be elicited.

To identify the epitope(s) recognized, cDNA encoding P501S was fragmented by various restriction digests, and sub-cloned into the retroviral expression vector pBIB-KS. Retroviral supernatants were generated by transfection of the helper packaging line Phoenix-Ampho. Supernatants were then used to transduce Jurkat/A2Kb cells for CTL screening. CTL were screened in IFN-gamma ELISPOT assays against these A2Kb targets transduced with the “library” of P501S fragments. Initial positive fragments P501S/H3 and P501S/F2 were sequenced and found to encode amino acids 106-553 and amino acids 136-547, respectively, of SEQ ID NO: 113. A truncation of H3 was made to encode amino acid residues 106-351 of SEQ ID NO: 113, which was unable to stimulate the CTL, thus localizing the epitope to amino acid residues 351-547. Additional fragments encoding amino acids 1-472 (Fragment A) and amino acids 1-351 (Fragment B) were also constructed. Fragment A but not Fragment B stimulated the CTL thus localizing the epitope to amino acid residues 351-472. Overlapping 20-mer and 18-mer peptides representing this region were tested by pulsing Jurkat/A2Kb cells versus CTL in an IFN-gamma assay. Only peptides P501S-369(20) and P501S-369(18) stimulated the CTL. Nine-mer and 10-mer peptides representing this region were synthesized and similarly tested. Peptide P501S-370 (SEQ ID NO: 539) was the minimal 9-mer giving a strong response. Peptide P501S-376 (SEQ ID NO: 540) also gave a weak response, suggesting that it might represent a cross-reactive epitope.

In subsequent studies, the ability of primary human B cells transduced with P501S to prime MHC class I-restricted, P501S-specific, autologous CD8 T cells was examined. Primary B cells were derived from PBMC of a homozygous HLA-A2 donor by culture in CD40 ligand and IL-4, transduced at high frequency with recombinant P501S in the vector pBIB, and selected with blastocidin-S. For *in vitro* priming, purified CD8⁺ T cells were cultured with autologous CD40 ligand + IL-4 derived, P501S-transduced B cells in a 96-well microculture format. These CTL microcultures were re-stimulated with P501S-transduced B cells and then assayed for specificity. Following this initial screen, microcultures with significant signal above background were cloned on autologous EBV-transformed B cells (BLCL), also transduced with P501S. Using IFN-gamma ELISPOT

for detection, several of these CD8 T cell clones were found to be specific for P501S, as demonstrated by reactivity to BLCL/P501S but not BLCL transduced with control antigen. It was further demonstrated that the anti-P501S CD8 T cell specificity is HLA-A2-restricted. First, antibody blocking experiments with anti-HLA-A,B,C monoclonal antibody (W6.32), anti-HLA-B,C monoclonal antibody (B1.23.2) and a control monoclonal antibody showed that only the anti-HLA-A,B,C antibody blocked recognition of P501S-expressing autologous BLCL. Secondly, the anti-P501S CTL also recognized an HLA-A2 matched, heterologous BLCL transduced with P501S, but not the corresponding EGFP transduced control BLCL.

A naturally processed, CD8, class I-restricted peptide epitope of P501S was identified as follows. Dendritic Cells (DC) were isolated by Percol gradient followed by differential adherence, and cultured for 5 days in the presence of RPMI medium containing 1% human serum, 50ng/ml GM-CSF and 30ng/ml IL-4. Following culture, DC were infected for 24 hours with P501S-expressing adenovirus at an MOI of 10 and matured for an additional 24 hours by the addition of 2ug/ml CD40 ligand. CD8 cells were enriched for by the subtraction of CD4+, CD14+ and CD16+ populations from PBMC with magnetic beads. Priming cultures containing 10,000 P501S-expressing DC and 100,000 CD8+ T cells per well were set up in 96-well V-bottom plates with RPMI containing 10% human serum, 5ng/ml IL-12 and 10ng/ml IL-6. Cultures were stimulated every 7 days using autologous fibroblasts retrovirally transduced to express P501S and CD80, and were treated with IFN-gamma for 48-72 hours to upregulate MHC Class I expression. 10u/ml IL-2 was added at the time of stimulation and on days 2 and 5 following stimulation. Following 4 stimulation cycles, one P501S-specific CD8+ T cell line (referred to as 2A2) was identified that produced IFN-gamma in response to IFN-gamma-treated P501S/CD80 expressing autologous fibroblasts, but not in response to IFN-gamma-treated P703P/CD80 expressing autologous fibroblasts in a γ -IFN Elispot assay. Line 2A2 was cloned in 96-well plates with 0.5 cell/well or 2 cells/well in the presence of 75,000 PBMC/well, 10,000 B-LCL/well, 30ng/ml OKT3 and 50u/ml IL-2. Twelve clones were isolated that showed strong P501S specificity in response to transduced fibroblasts.

Fluorescence activated cell sorting (FACS) analysis was performed on P501S-specific clones using CD3-, CD4- and CD8-specific antibodies conjugated to PercP, FITC and PE respectively. Consistent with the use of CD8 enriched T cells in the priming cultures, P5401S-specific clones were determined to be CD3+, CD8+ and CD4-.

5 To identify the relevant P501S epitope recognized by P501S specific CTL, pools of 18-20 mer or 30-mer peptides that spanned the majority of the amino acid sequence of P501S were loaded onto autologous B-LCL and tested in γ -IFN Elispot assays for the ability to stimulate two P501S-specific CTL clones, referred to as 4E5 and 4E7. One pool, composed of five 18-20 mer peptides that spanned amino acids 411-486 of P501S
10 (SEQ ID NO: 113), was found to be recognized by both P501S-specific clones. To identify the specific 18-20 mer peptide recognized by the clones, each of the 18-20 mer peptides that comprised the positive pool were tested individually in γ -IFN Elispot assays for the ability to stimulate the two P501S-specific CTL clones, 4E5 and 4E7. Both 4E5 and 4E7 specifically recognized one 20-mer peptide (SEQ ID NO: 853; cDNA sequence provided in
15 SEQ ID NO: 854) that spanned amino acids 453-472 of P501S. Since the minimal epitope recognized by CD8+ T cells is almost always either a 9 or 10-mer peptide sequence, 10-mer peptides that spanned the entire sequence of SEQ ID NO: 853 were synthesized that differed by 1 amino acid. Each of these 10-mer peptides was tested for the ability to stimulate two P501S-specific clones, (referred to as 1D5 and 1E12). One 10-mer peptide
20 (SEQ ID NO: 855; cDNA sequence provided in SEQ ID NO: 856) was identified that specifically stimulated the P501S-specific clones. This epitope spans amino acids 463-472 of P501S. This sequence defines a minimal 10-mer epitope from P501S that can be naturally processed and to which CTL responses can be identified in normal PBMC. Thus, this epitope is a candidate for use as a vaccine moiety, and as a therapeutic and/or
25 diagnostic reagent for prostate cancer.

To identify the class I restriction element for the P501S-derived sequence of SEQ ID NO: 855, HLA blocking and mismatch analyses were performed. In γ -IFN Elispot assays, the specific response of clones 4A7 and 4E5 to P501S-transduced autologous fibroblasts was blocked by pre-incubation with 25ug/ml W6/32 (pan-Class I blocking

antibody) and B1.23.2 (HLA-B/C blocking antibody). These results demonstrate that the SEQ ID NO: 855-specific response is restricted to an HLA-B or HLA-C allele.

For the HLA mismatch analysis, autologous B-LCL (HLA-A1,A2,B8,B51, Cw1, Cw7) and heterologous B-LCL (HLA-A2,A3,B18,B51,Cw5,Cw14) that share the HLAB51 allele were pulsed for one hour with 20ug/ml of peptide of SEQ ID NO: 855, washed, and tested in γ -IFN Elispot assays for the ability to stimulate clones 4A7 and 4E5. Antibody blocking assays with the B1.23.2 (HLA-B/C blocking antibody) were also performed. SEQ ID NO: 855-specific response was detected using both the autologous (D326) and heterologous (D107) B-LCL, and furthermore the responses were blocked by pre-incubation with 25ug/ml of B1.23.2 HLA-B/C blocking antibody. Together these results demonstrate that the P501S-specific response to the peptide of SEQ ID NO: 855 is restricted to the HLA-B51 class I allele. Molecular cloning and sequence analysis of the HLA-B51 allele from D3326 revealed that the HLA-B51 subtype of D326 is HLA-B51011.

Based on the 10-mer P501S-derived epitope of SEQ ID NO: 855, two 9-mers with the sequences of SEQ ID NO: 857 and 858 were synthesized and tested in Elispot assays for the ability to stimulate two P501S-specific CTL clones derived from line 2A2. The 10-mer peptide of SEQ ID NO: 855, as well as the 9-mer peptide of SEQ ID NO: 858, but not the 9-mer peptide of SEQ ID NO: 857, were capable of stimulating the P501S-specific CTL to produce IFN-gamma. These results demonstrate that the peptide of SEQ ID NO: 858 is a 9-mer P501S-derived epitope recognized by P501S-specific CTL. The DNA sequence encoding the epitope of SEQ ID NO: 858 is provided in SEQ ID NO: 859.

To identify the class I restricting allele for the P501S-derived peptide of SEQ ID NO: 855 and 858 specific response, each of the HLA B and C alleles were cloned from the donor used in the *in vitro* priming experiment. Sequence analysis indicated that the relevant alleles were HLA-B8, HLA-B51, HLA-Cw01 and HLA-Cw07. Each of these alleles were subcloned into an expression vector and co-transfected together with the P501S gene into VA-13 cells. Transfected VA-13 cells were then tested for the ability to specifically stimulate the P501S-specific CTL in ELISPOT assays. VA-13 cells

transfected with P501S and HLA-B51 were capable of stimulating the P501S-specific CTL to secrete gamma-IFN. VA-13 cells transfected with HLA-B51 alone or P501S + the other HLA-alleles were not capable of stimulating the P501S-specific CTL. These results demonstrate that the restricting allele for the P501S-specific response is the HLAB51
 5 allele. Sequence analysis revealed that the subtype of the relevant restricting allele is HLA-B51011.

A naturally processed CD4 epitope of P501S was identified as follows.

CD4 cells specific for P501S were prepared as described above. A series of 16 overlapping peptides were synthesized that spanned approximately 50% of the amino
 10 terminal portion of the P501S gene (amino acids 1- 325 of SEQ ID NO: 113). For priming, peptides were combined into pools of 4 peptides, pulsed at 4 µg/ml onto dendritic cells (DC) for 24 hours, with TNF-alpha. DC were then washed and mixed with negatively selected CD4+ T cells in 96 well U-bottom plates. Cultures were re-stimulated weekly on
 15 fresh DC loaded with peptide pools. Following a total of 4 stimulation cycles, cells were rested for an additional week and tested for specificity to APC pulsed with peptide pools using γ-IFN ELISA and proliferation assays. For these assays, adherent monocytes loaded with either the relevant peptide pool at 4ug/ml or an irrelevant peptide at µg/ml were used as APC. T cell lines that demonstrated either specific cytokine secretion or proliferation were then tested for recognition of individual peptides that were present in the pool. T cell
 20 lines could be identified from pools A and B that recognized individual peptides from these pools.

From pool A, lines AD9 and AE10 specifically recognized peptide 1 (SEQ ID NO: 862), and line AF5 recognized peptide 39 (SEQ ID NO: 861). From pool B, line BC6 could be identified that recognized peptide 58 (SEQ ID NO: 860). Each of these lines were stimulated on the specific peptide and tested for specific recognition of the peptide in a titration assay as well as cell lysates generated by infection of HEK 293 cells with adenovirus expressing either P501S or an irrelevant antigen. For these assays, APC-adherent monocytes were pulsed with either 10, 1, or 0.1 µg/ml individual P501S peptides, and DC were pulsed overnight with a 1:5 dilution of adenovirally infected cell lysates. Lines AD9, AE10 and AF5 retained significant recognition of the relevant P501S-derived peptides even at 0.1 mg/ml. Furthermore, line Ad9 demonstrated significant (8.1 fold stimulation index) specific activity for lysates from adenovirus-P501S infected cells. These results demonstrate that high affinity CD4 T cell lines can be generated toward P501S-derived epitopes, and that at least a subset of these T cells specific for the P501S derived sequence of SEQ ID NO: 862 are specific for an epitope that is naturally processed by human cells. The DNA sequences encoding the amino acid sequences of SEQ ID NO: 860-862 are provided in SEQ ID NO: 863-865, respectively.

EXAMPLE 13

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS

BY MICROARRAY ANALYSIS

This Example describes the isolation of certain prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 372 clones were identified, and 319 were successfully sequenced. Table I presents a summary of these clones, which are shown in SEQ ID NOs:385-400. Of these sequences SEQ ID NOs:386, 389, 390 and 392 correspond to

Summary of Prostate Tumor Antigens

Known Genes	Previously Identified Genes	Novel Genes
T-cell gamma chain	P504S	23379 (SEQ ID NO:389)
Kallikrein	P1000C	23399 (SEQ ID NO:392)
Vector	P501S	23320 (SEQ ID NO:386)
CGI-82 protein mRNA (23319; SEQ ID NO:385)	P503S	23381 (SEQ ID NO:390)
PSA	P510S	
Ald. 6 Dehyd.	P784P	
L-iditol-2 dehydrogenase (23376; SEQ ID NO:388)	P502S	
Ets transcription factor PDEF (22672; SEQ ID NO:398)	P706P	
hTGR (22678; SEQ ID NO:399)	19142.2, bangur.seq (22621; SEQ ID NO:396)	
KIAA0295(22685; SEQ ID NO:400)	5566.1 Wang (23404; SEQ ID NO:393)	
Prostatic Acid Phosphatase(22655; SEQ ID NO:397)	P712P	
transglutaminase (22611; SEQ ID NO:395)	P778P	
HDLBP (23508; SEQ ID NO:394)		
CGI-69 Protein(23367; SEQ ID NO:387)		
KIAA0122(23383; SEQ ID NO:391)		
TEEG		

CGI-82 showed 4.06 fold over-expression in prostate tissues as compared to
 5 other normal tissues tested. It was over-expressed in 43% of prostate tumors, 25% normal

prostate, not detected in other normal tissues tested. L-iditol-2 dehydrogenase showed 4.94 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 90% of prostate tumors, 100% of normal prostate, and not detected in other normal tissues tested. Ets transcription factor PDEF showed 5.55 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% prostate tumors, 25% normal prostate and not detected in other normal tissues tested. hTGR1 showed 9.11 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 63% of prostate tumors and is not detected in normal tissues tested including normal prostate. KIAA0295 showed 5.59 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% of prostate tumors, low to undetectable in normal tissues tested including normal prostate tissues. Prostatic acid phosphatase showed 9.14 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 67% of prostate tumors, 50% of normal prostate, and not detected in other normal tissues tested. Transglutaminase showed 14.84 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 30% of prostate tumors, 50% of normal prostate, and is not detected in other normal tissues tested. High density lipoprotein binding protein (HDLBP) showed 28.06 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% of normal prostate, and is undetectable in all other normal tissues tested. CGI-69 showed 3.56 fold over-expression in prostate tissues as compared to other normal tissues tested. It is a low abundant gene, detected in more than 90% of prostate tumors, and in 75% normal prostate tissues. The expression of this gene in normal tissues was very low. KIAA0122 showed 4.24 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 57% of prostate tumors, it was undetectable in all normal tissues tested including normal prostate tissues. 19142.2 bangur showed 23.25 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors and 100% of normal prostate. It was undetectable in other normal tissues tested. 5566.1 Wang showed 3.31 fold over-

expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% normal prostate and was also over-expressed in normal bone marrow, pancreas, and activated PBMC. Novel clone 23379 (also referred to as P553S) showed 4.86 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in 97% of prostate tumors and 75% normal prostate and is undetectable in all other normal tissues tested. Novel clone 23399 showed 4.09 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 27% of prostate tumors and was undetectable in all normal tissues tested including normal prostate tissues. Novel clone 23320 showed 3.15 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in all prostate tumors and 50% of normal prostate tissues. It was also expressed in normal colon and trachea. Other normal tissues do not express this gene at high level.

Subsequent full-length cloning studies on P553S, using standard techniques, revealed that this clone is an incomplete spliced form of P501S. The determined cDNA sequences for four splice variants of P553S are provided in SEQ ID NO: 702-705. An amino acid sequence encoded by SEQ ID NO: 705 is provided in SEQ ID NO: 706. The cDNA sequence of SEQ ID NO: 702 was found to contain two open reading frames (ORFs). The amino acid sequences encoded by these two ORFs are provided in SEQ ID NO: 707 and 708.

EXAMPLE 14

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS

BY ELECTRONIC SUBTRACTION

This Example describes the use of an electronic subtraction technique to identify prostate-specific antigens.

Potential prostate-specific genes present in the GenBank human EST database were identified by electronic subtraction (similar to that described by Vasmatizis et al., *Proc. Natl. Acad. Sci. USA* 95:300-304, 1998). The sequences of EST clones

(43,482) derived from various prostate libraries were obtained from the GenBank public human EST database. Each prostate EST sequence was used as a query sequence in a BLASTN (National Center for Biotechnology Information) search against the human EST database. All matches considered identical (length of matching sequence >100 base pairs, density of identical matches over this region > 70%) were grouped (aligned) together in a cluster. Clusters containing more than 200 ESTs were discarded since they probably represented repetitive elements or highly expressed genes such as those for ribosomal proteins. If two or more clusters shared common ESTs, those clusters were grouped together into a "supercluster," resulting in 4,345 prostate superclusters.

Records for the 479 human cDNA libraries represented in the GenBank release were downloaded to create a database of these cDNA library records. These 479 cDNA libraries were grouped into three groups: Plus (normal prostate and prostate tumor libraries, and breast cell line libraries, in which expression was desired), Minus (libraries from other normal adult tissues, in which expression was not desirable), and Other (libraries from fetal tissue, infant tissue, tissues found only in women, non-prostate tumors and cell lines other than prostate cell lines, in which expression was considered to be irrelevant). A summary of these library groups is presented in Table II.

Table II

Prostate cDNA Libraries and ESTs

Library	# of Libraries	# of ESTs
Plus	25	43,482
Normal	11	18,875
Tumor	11	21,769
Cell lines	3	2,838
Minus	166	
Other	287	

Each supercluster was analyzed in terms of the ESTs within the supercluster. The tissue source of each EST clone was noted and used to classify the superclusters into four groups: Type 1- EST clones found in the Plus group libraries only; no expression detected in Minus or Other group libraries; Type 2- EST clones derived from the Plus and Other group libraries only; no expression detected in the Minus group; Type 3- EST clones derived from the Plus, Minus and Other group libraries, but the number of ESTs derived from the Plus group is higher than in either the Minus or Other groups; and Type 4- EST clones derived from Plus, Minus and Other group libraries, but the number derived from the Plus group is higher than the number derived from the Minus group. This analysis identified 4,345 breast clusters (*see* Table III). From these clusters, 3,172 EST clones were ordered from Research Genetics, Inc., and were received as frozen glycerol stocks in 96-well plates.

Table III

Prostate Cluster Summary

Type	# of Superclusters	# of ESTs Ordered
1	688	677
2	2899	2484
3	85	11
4	673	0
Total	4345	3172

The EST clone inserts were PCR-amplified using amino-linked PCR primers for Synteni microarray analysis. When more than one PCR product was obtained for a particular clone, that PCR product was not used for expression analysis. In total, 2,528 clones from the electronic subtraction method were analyzed by microarray analysis to identify electronic subtraction breast clones that had high levels of tumor vs. normal

tissue mRNA. Such screens were performed using a Synteni (Palo Alto, CA) microarray, according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Within these analyses, the clones were arrayed on the chip, which was then probed with fluorescent probes generated from normal and tumor prostate cDNA, as well as various other normal tissues. The slides were scanned and the fluorescence intensity was measured.

Clones with an expression ratio greater than 3 (*i.e.*, the level in prostate tumor and normal prostate mRNA was at least three times the level in other normal tissue mRNA) were identified as prostate tumor-specific sequences (Table IV). The sequences of these clones are provided in SEQ ID NO: 401-453, with certain novel sequences shown in SEQ ID NO: 407, 413, 416-419, 422, 426, 427 and 450.

Table IV

Prostate-tumor Specific Clones

SEQ ID NO.	Sequence Designation	Comments
401	22545	previously identified P1000C
402	22547	previously identified P704P
403	22548	known
404	22550	known
405	22551	PSA
406	22552	prostate secretory protein 94
407	22553	novel
408	22558	previously identified P509S
409	22562	glandular kallikrein
410	22565	previously identified P1000C
411	22567	PAP
412	22568	B1006C (breast tumor antigen)
413	22570	novel
414	22571	PSA
415	22572	previously identified P706P
416	22573	novel

417	22574	novel
418	22575	novel
419	22580	novel
420	22581	PAP
421	22582	prostatic secretory protein 94
422	22583	novel
423	22584	prostatic secretory protein 94
424	22585	prostatic secretory protein 94
425	22586	known
426	22587	novel
427	22588	novel
428	22589	PAP
429	22590	known
430	22591	PSA
431	22592	known
432	22593	Previously identified P777P
433	22594	T cell receptor gamma chain
434	22595	Previously identified P705P
435	22596	Previously identified P707P
436	22847	PAP
437	22848	known
438	22849	prostatic secretory protein 57
439	22851	PAP
440	22852	PAP
441	22853	PAP
442	22854	previously identified P509S
443	22855	previously identified P705P
444	22856	previously identified P774P
445	22857	PSA
446	23601	previously identified P777P
447	23602	PSA
448	23605	PSA
449	23606	PSA
450	23612	novel
451	23614	PSA
452	23618	previously identified P1000C
453	23622	previously identified P705P

Further studies on the clone of SEQ ID NO: 407 (also referred to as P1020C) led to the isolation of an extended cDNA sequence provided in SEQ ID NO: 591.

This extended cDNA sequence was found to contain an open reading frame that encodes the predicted amino acid sequence of SEQ ID NO: 592. The P1020C cDNA and amino acid sequences were found to show some similarity to the human endogenous retroviral HERV-K pol gene and protein.

5

EXAMPLE 15

FURTHER IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY MICROARRAY ANALYSIS

This Example describes the isolation of additional prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 142 clones were identified and sequenced. Certain of these clones are shown in SEQ ID NO: 454-467. Of these sequences, SEQ ID NO: 459-461 represent novel genes. The others (SEQ ID NO: 454-458 and 461-467) correspond to known sequences.

EXAMPLE 16

FURTHER CHARACTERIZATION OF PROSTATE-SPECIFIC ANTIGEN P710P

This Example describes the full length cloning of P710P.

The prostate cDNA library described above was screened with the P710P fragment described above. One million colonies were plated on LB/Ampicillin plates. Nylon membrane filters were used to lift these colonies, and the cDNAs picked up by these filters were then denatured and cross-linked to the filters by UV light. The P710P fragment was radiolabeled and used to hybridize with the filters. Positive cDNA clones were selected and their cDNAs recovered and sequenced by an automatic Perkin Elmer/Applied Biosystems Division Sequencer. Four sequences were obtained, and are presented in SEQ

ID NO: 468-471. These sequences appear to represent different splice variants of the P710P gene. Subsequent comparison of the cDNA sequences of P710P with those in Genbank revealed homology to the DD3 gene (Genbank accession numbers AF103907 & AF103908). The cDNA sequence of DD3 is provided in SEQ ID NO: 690.

5

EXAMPLE 17

PROTEIN EXPRESSION OF PROSTATE-SPECIFIC ANTIGENS

This example describes the expression and purification of prostate-specific
10 antigens in *E. coli*, baculovirus and mammalian cells.

a) Expression of P501S in *E. coli*

Expression of the full-length form of P501S was attempted by first cloning P501S without the leader sequence (amino acids 36-553 of SEQ ID NO: 113) downstream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) in
15 pET17b. Specifically, P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW003 (SEQ ID NO: 486). AW025 is a sense cloning primer that contains a HindIII site. AW003 is an antisense cloning primer that contains an EcoRI site. DNA amplification was performed using 5 µl 10X Pfu buffer, 1 µl 20 mM dNTPs, 1 µl each of the PCR primers at 10 µM concentration, 40 µl water, 1 µl Pfu DNA polymerase
20 (Stratagene, La Jolla, CA) and 1 µl DNA at 100 ng/µl. Denaturation at 95°C was performed for 30 sec, followed by 10 cycles of 95°C for 30 sec, 60°C for 1 min and by 72°C for 3 min. 20 cycles of 95°C for 30 sec, 65°C for 1 min and by 72°C for 3 min, and lastly by 1 cycle of 72°C for 10 min. The PCR product was cloned to Ra12m/pET17b using HindIII and EcoRI. The sequence of the resulting fusion construct (referred to as
25 Ra12-P501S-F) was confirmed by DNA sequencing.

The fusion construct was transformed into BL21(DE3)pLysE, pLysS and CodonPlus *E. coli* (Stratagene) and grown overnight in LB broth with kanamycin. The resulting culture was induced with IPTG. Protein was transferred to PVDF membrane and

blocked with 5% non-fat milk (in PBS-Tween buffer), washed three times and incubated with mouse anti-His tag antibody (Clontech) for 1 hour. The membrane was washed 3 times and probed with HRP-Protein A (Zymed) for 30 min. Finally, the membrane was washed 3 times and developed with ECL (Amersham). No expression was detected by Western blot. Similarly, no expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage (Invitrogen).

An N-terminal fragment of P501S (amino acids 36-325 of SEQ ID NO: 113) was cloned down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 in pET17b as follows. P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW027 (SEQ ID NO: 487). AW027 is an antisense cloning primer that contains an EcoRI site and a stop codon. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The fusion construct (referred to as Ra12-P501S-N) was confirmed by DNA sequencing.

The Ra12-P501S-N fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, essentially as described above. Using Western blot analysis, protein bands were observed at the expected molecular weight of 36 kDa. Some high molecular weight bands were also observed, probably due to aggregation of the recombinant protein. No expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage.

A fusion construct comprising a C-terminal portion of P501S (amino acids 257-553 of SEQ ID NO: 113) located down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) was prepared as follows. P501S DNA was used to perform PCR using the primers AW026 (SEQ ID NO: 488) and AW003 (SEQ ID NO: 486). AW026 is a sense cloning primer that contains a HindIII site. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The sequence for the fusion construct (referred to as Ra12-P501S-C) was confirmed.

The Ra12-P501S-C fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, as described above. A small amount of protein was detected by Western blot, with some molecular weight aggregates also being observed. Expression was also detected by Western blot when the Ra12-P501S-C fusion was used for
 5 expression in BL21CodonPlus induced by CE6 phage.

A fusion construct comprising a fragment of P501S (amino acids 36-298 of SEQ ID NO: 113) located down-stream of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 848) was prepared as follows. P501S DNA was used to perform PCR using the primers AW042 (SEQ ID NO: 849) and AW053 (SEQ ID NO: 850). AW042 is a sense cloning
 10 primer that contains a EcoRI site. AW053 is an antisense primer with stop and Xho I sites. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the EcoRI and Xho I sites. The resulting fusion construct (referred to as Ra12-P501S-E2) was expressed in B834 (DE3) pLys S *E. coli* host cells in TB media for 2 h at room temperature. Expressed protein was purified by washing
 15 the inclusion bodies and running on a Ni-NTA column. The purified protein stayed soluble in buffer containing 20 mM Tris-HCl (pH 8), 100 mM NaCl, 10 mM β -Me and 5% glycerol. The determined cDNA and amino acid sequences for the expressed fusion protein are provided in SEQ ID NO: 851 and 852, respectfully.

20 b) Expression of P501S in Baculovirus

The Bac-to-Bac baculovirus expression system (BRL Life Technologies, Inc.) was used to express P501S protein in insect cells. Full-length P501S (SEQ ID NO: 113) was amplified by PCR and cloned into the XbaI site of the donor plasmid pFastBacI. The recombinant bacmid and baculovirus were prepared according to the manufacturer's
 25 instructions. The recombinant baculovirus was amplified in Sf9 cells and the high titer viral stocks were utilized to infect High Five cells (Invitrogen) to make the recombinant protein. The identity of the full-length protein was confirmed by N-terminal sequencing of the recombinant protein and by Western blot analysis (Figure 7). Specifically, 0.6 million

High Five cells in 6-well plates were infected with either the unrelated control virus BV/ECD_PD (lane 2), with recombinant baculovirus for P501S at different amounts or MOIs (lanes 4-8), or were uninfected (lane 3). Cell lysates were run on SDS-PAGE under reducing conditions and analyzed by Western blot with the anti-P501S monoclonal antibody P501S-10E3-G4D3 (prepared as described below). Lane 1 is the biotinylated protein molecular weight marker (BioLabs).

The localization of recombinant P501S in the insect cells was investigated as follows. The insect cells overexpressing P501S were fractionated into fractions of nucleus, mitochondria, membrane and cytosol. Equal amounts of protein from each fraction were analyzed by Western blot with a monoclonal antibody against P501S. Due to the scheme of fractionation, both nucleus and mitochondria fractions contain some plasma membrane components. However, the membrane fraction is basically free from mitochondria and nucleus. P501S was found to be present in all fractions that contain the membrane component, suggesting that P501S may be associated with plasma membrane of the insect cells expressing the recombinant protein.

c) Expression of P501S in mammalian cells

Full-length P501S (553AA) was cloned into various mammalian expression vectors, including pCEP4 (Invitrogen), pVR1012 (Vical, San Diego, CA) and a modified form of the retroviral vector pBMN, referred to as pBIB. Transfection of P501S/pCEP4 and P501S/pVR1012 into HEK293 fibroblasts was carried out using the Fugene transfection reagent (Boehringer Mannheim). Briefly, 2 ul of Fugene reagent was diluted into 100 ul of serum-free media and incubated at room temperature for 5-10 min. This mixture was added to 1 ug of P501S plasmid DNA, mixed briefly and incubated for 30 minutes at room temperature. The Fugene/DNA mixture was added to cells and incubated for 24-48 hours. Expression of recombinant P501S in transfected HEK293 fibroblasts was detected by means of Western blot employing a monoclonal antibody to P501S.

Transfection of p501S/pCEP4 into CHO-K cells (American Type Culture Collection, Rockville, MD) was carried out using GenePorter transfection reagent (Gene

Therapy Systems, San Diego, CA). Briefly, 15 µl of GenePorter was diluted in 500 µl of serum-free media and incubated at room temperature for 10 min. The GenePorter/media mixture was added to 2 µg of plasmid DNA that was diluted in 500 µl of serum-free media, mixed briefly and incubated for 30 min at room temperature. CHO-K cells were rinsed in
 5 PBS to remove serum proteins, and the GenePorter/DNA mix was added and incubated for 5 hours. The transfected cells were then fed an equal volume of 2x media and incubated for 24-48 hours.

FACS analysis of P501S transiently infected CHO-K cells, demonstrated surface expression of P501S. Expression was detected using rabbit polyclonal antisera
 10 raised against a P501S peptide, as described below. Flow cytometric analysis was performed using a FaCScan (Becton Dickinson), and the data were analyzed using the Cell Quest program.

d) Expression of P703P in Baculovirus

The cDNA for full-length P703P-DE5 (SEQ ID NO: 326), together with
 15 several flanking restriction sites, was obtained by digesting the plasmid pCDNA703 with restriction endonucleases Xba I and Hind III. The resulting restriction fragment (approx. 800 base pairs) was ligated into the transfer plasmid pFastBacI which was digested with the same restriction enzymes. The sequence of the insert was confirmed by DNA sequencing. The recombinant transfer plasmid pFBP703 was used to make recombinant bacmid DNA
 20 and baculovirus using the Bac-To-Bac Baculovirus expression system (BRL Life Technologies). High Five cells were infected with the recombinant virus BVP703, as described above, to obtain recombinant P703P protein.

e) Expression of P788P in *E. Coli*

A truncated, N-terminal portion, of P788P (residues 1-644 of SEQ ID NO:
 25 777; referred to as P788P-N) fused with a C-terminal 6xHis Tag was expressed in *E. coli* as follows. P788P cDNA was amplified using the primers AW080 and AW081 (SEQ ID NO: 815 and 816). AW080 is a sense cloning primer with an NdeI site. AW081 is an antisense

cloning primer with a XhoI site. The PCR-amplified P788P, as well as the vector pCRX1, were digested with NdeI and XhoI. Vector and insert were ligated and transformed into NovaBlue cells. Colonies were randomly screened for insert and then sequenced. P788P-N clone #6 was confirmed to be identical to the designed construct. The expression
 5 construct P788P-N #6/pCRX1 was transformed into *E. coli* BL21 CodonPlus-RIL competent cells. After induction, most of the cells grew well, achieving OD600 of greater than 2.0 after 3 hr. Coomassie stained SDS-PAGE showed an over-expressed band at about 75 kD. Western blot analysis using a 6xHisTag antibody confirmed the band was P788P-N. The determined cDNA sequence for P788P-N is provided in SEQ ID NO: 817,
 10 with the corresponding amino acid sequence being provided in SEQ ID NO: 818.

f) Expression of P510S in *E. Coli*

The P510S protein has 9 potential transmembrane domains and is predicted to be located at the plasma membrane. The C-terminal protein of this protein, as well as the predicted third extracellular domain of P510S were expressed in *E. coli* as follows.

15 The expression construct referred to as Ra12-P501S-C was designed to have a 6 HisTag at the N-terminal end, followed by the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 819) and then the C-terminal portion of P510S (amino residues 1176-1261 of SEQ ID NO: 538). Full-length P510S was used to amplify the P510S-C fragment by PCR using the primers AW056 and AW057 (SEQ ID NO: 820 and 821, respectively). AW056 is a sense
 20 cloning primer with an EcoRI site. AW057 is an antisense primer with stop and XhoI sites. The amplified P501S fragment and Ra12/pCRX1 were digested with EcoRI and XhoI and then purified. The insert and vector were ligated together and transformed into NovaBlue. Colonies were randomly screened for insert and sequences. For protein expression, the expression construct was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent
 25 cells. A mini-induction screen was performed to optimize the expression conditions. After induction the cells grew well, achieving OD 600 nm greater than 2.0 after 3 hours. Coomassie stain SDS-PAGE showed a highly over-expressed band at approx. 30 kD. Though this is higher than the expected molecular weight, western blot analysis was

positive, showing this band to be the His tag-containing protein. The optimized culture conditions are as follows. Dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2xYT (with kanamycin and chloramphenicol) at a ratio of 25 ml culture to 1 liter 2xYT. Allow to grow at 37 °C until OD600 = 0.6. Take an aliquot out as
 5 T0 sample. Add 1 mM IPTG and allow to grow at 30 °C for 3 hours. Take out a T3 sample, spin down cells and store at -80 °C. The determined cDNA and amino acid sequences for the Ra12-P510S-C construct are provided in SEQ ID NO: 822 and 825, respectively.

The expression construct P510S-C was designed to have a 5' added start codon and
 10 a glycine (GGA) codon and then the P510S C terminal fragment followed by the in frame 6x histidine tag and stop codon from the pET28b vector. The cloning strategy is similar to that used for Ra12-P510S-C, except that the PCR primers employed were those shown in SEQ ID NO: 828 and 829, respectively and the NcoI/XhoI cut in pET28b was used. The primer of SEQ ID NO: 828 created a 5' NcoI site and added a start codon. The antisense
 15 primer of SEQ ID NO: 829 creates a XhoI site on P510S C terminal fragment. Clones were confirmed by sequencing. For protein expression, the expression construct was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. An OD600 of greater than 2.0 was obtained 30 hours after induction. Coomassie stained SDS-PAGE showed an over-expressed band at about 11 kD. Western blot analysis confirmed that the
 20 band was P510S-C, as did N-terminal protein sequencing. The optimized culture conditions are as follows: dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2x YT (+ kanamycin and chloramphenicol) at a ratio of 25 mL culture to 1 liter 2x YT, and allow to grow at 37 °C until an OD 600 of about 0.5 is reached. Take out an aliquot as T0 sample. Add 1 mM IPTG and allow to grow at 30 °C for 3 hours.
 25 Spin down the cells and store at -80 °C until purification. The determined cDNA and amino acid sequence for the P510S-C construct are shown in SEQ ID NO: 823 and 826, respectively.

The predicted third extracellular domain of P510S (P510S-E3; residues 328-676 of SEQ ID NO: 538) was expressed in *E. coli* as follows. The P510S fragment was

amplified by PCR using the primers shown in SEQ ID NO: 830 and 831. The primer of SEQ ID NO: 830 is a sense primer with an NdeI site for use in ligating into pPDM. The primer of SEQ ID NO: 831 is an antisense primer with an added XhoI site for use in ligating into pPDM. The resulting fragment was cloned to pPDM at the NdeI and XhoI sites. Clones were confirmed by sequencing. For protein expression, the clone was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. After induction, an OD600 of greater than 2.0 was achieved after 3 hours. Coomassie stained SDS-PAGE showed an over-expressed band at about 39 kD, and N-terminal sequencing confirmed the N-terminal to be that of P510S-E3. Optimized culture conditions are as follows: dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2x YT (kanamycin and chloramphenicol) at a ratio of 25 ml culture to 1 liter 2x YT. Allow to grow at 37 °C until OD 600 equals 0.6. Take out an aliquot as T0 sample. Add 1 mM IPTG and allow to grow at 30 °C for 3 hours. Take out a T3 sample, spin down the cells and store at -80 °C until purification. The determined cDNA and amino acid sequences for the P501S-E3 construct are provided in SEQ ID NO: 824 and 827, respectively.

g) Expression of P775S in *E. Coli*

The antigen P775P contains multiple open reading frames (ORF). The third ORF, encoding the protein of SEQ ID NO: 483, has the best motif score. An expression fusion construct containing the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 819) and P775P-ORF3 with an N-terminal 6x HisTag was prepared as follows. P775P-ORF3 was amplified using the sense PCR primers of SEQ ID NO: 832 and the anti-sense PCR primer of SEQ ID NO: 833. The PCR amplified fragment of P775P and Ra12/pCRX1 were digested with the restriction enzymes EcoRI and XhoI. Vector and insert were ligated and then transformed into NovaBlue cells. Colonies were randomly screened for insert and then sequenced. A clone having the desired sequence was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. Two hours after induction, the cell density peaked at OD600 of approximately 1.8. Coomassie stained SDS-PAGE showed an over-expressed band at about 31 kD. Western blot using 6x HisTag antibody confirmed that the band was

Ra12-P775P-ORF3. The determined cDNA and amino acid sequences for the fusion construct are provided in SEQ ID NO: 834 and 835, respectively.

H) EXPRESSION OF A P703P HIS TAG FUSION PROTEIN IN E. COLI

5 The cDNA for the coding region of P703P was prepared by PCR using the primers of SEQ ID NO: 836 and 837. The PCR product was digested with EcoRI restriction enzyme, gel purified and cloned into a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into *E. coli*
10 BL21 (DE3) pLys S expression host cells. The determined amino acid and cDNA sequences for the expressed recombinant P703P are provided in SEQ ID NO: 838 and 839, respectively.

I) EXPRESSION OF A P705P HIS TAG FUSION PROTEIN IN E. COLI

15 The cDNA for the coding region of P705P was prepared by PCR using the primers of SEQ ID NO: 840 and 841. The PCR product was digested with EcoRI restriction enzyme, gel purified and cloned into a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into *E. coli*
20 BL21 (DE3) pLys S and BL21 (DE3) CodonPlus expression host cells. The determined amino acid and cDNA sequences for the expressed recombinant P705P are provided in SEQ ID NO: 842 and 843, respectively.

J) EXPRESSION OF A P711P HIS TAG FUSION PROTEIN IN *E. COLI*

The cDNA for the coding region of P711P was prepared by PCR using the primers of SEQ ID NO: 844 and 845. The PCR product was digested with EcoRI restriction enzyme, gel purified and cloned into a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into *E. coli* BL21 (DE3) pLys S and BL21 (DE3) CodonPlus expression host cells. The determined amino acid and cDNA sequences for the expressed recombinant P711P are provided in SEQ ID NO: 846 and 847, respectively.

EXAMPLE 18

PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST PROSTATE-SPECIFIC POLYPEPTIDES

a) Preparation and Characterization of Polyclonal Antibodies against P703P, P504S and P509S

Polyclonal antibodies against P703P, P504S and P509S were prepared as follows.

Each prostate tumor antigen expressed in an *E. coli* recombinant expression system was grown overnight in LB broth with the appropriate antibiotics at 37°C in a shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml to 2x YT plus appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When the Optical Density (at 560 nm) of the culture reached 0.4-0.6, the cells were induced with IPTG (1 mM). Four hours after induction with IPTG, the cells were harvested by centrifugation. The cells were then washed with phosphate buffered saline and centrifuged again. The supernatant was discarded and the cells were either frozen for future use or immediately processed. Twenty ml of lysis buffer was added to the cell pellets and vortexed. To break

open the *E. coli* cells, this mixture was then run through the French Press at a pressure of 16,000 psi. The cells were then centrifuged again and the supernatant and pellet were checked by SDS-PAGE for the partitioning of the recombinant protein. For proteins that localized to the cell pellet, the pellet was resuspended in 10 mM Tris pH 8.0, 1% CHAPS and the inclusion body pellet was washed and centrifuged again. This procedure was repeated twice more. The washed inclusion body pellet was solubilized with either 8 M urea or 6 M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The solubilized protein was added to 5 ml of nickel-chelate resin (Qiagen) and incubated for 45 min to 1 hour at room temperature with continuous agitation. After incubation, the resin and protein mixture were poured through a disposable column and the flow through was collected. The column was then washed with 10-20 column volumes of the solubilization buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin such as HiPrepQ (Biorad) was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Each antigen was eluted off the column with a increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool. The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The proteins were then vialled after filtration through a 0.22 micron filter and the antigens were frozen until needed for immunization.

Four hundred micrograms of each prostate antigen was combined with 100 micrograms of muramyldipeptide (MDP). Every four weeks rabbits were boosted with 100 micrograms mixed with an equal volume of Incomplete Freund's Adjuvant (IFA). Seven days following each boost, the animal was bled. Sera was generated by incubating the blood at 4°C for 12-4 hours followed by centrifugation.

Ninety-six well plates were coated with antigen by incubating with 50 microliters (typically 1 microgram) of recombinant protein at 4 °C for 20 hours. 250 microliters of BSA blocking buffer was added to the wells and incubated at room

temperature for 2 hours. Plates were washed 6 times with PBS/0.01% Tween. Rabbit sera was diluted in PBS. Fifty microliters of diluted sera was added to each well and incubated at room temperature for 30 min. Plates were washed as described above before 50 microliters of goat anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 min. Plates were again washed as described above and 100 microliters of TMB microwell peroxidase substrate was added to each well. Following a 15 min incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H₂SO₄ and read immediately at 450 nm. All polyclonal antibodies showed immunoreactivity to the appropriate antigen.

10 b) Preparation and Characterization of Antibodies against P501S

A murine monoclonal antibody directed against the carboxy-terminus of the prostate-specific antigen P501S was prepared as follows.

A truncated fragment of P501S (amino acids 355-526 of SEQ ID NO: 113) was generated and cloned into the pET28b vector (Novagen) and expressed in *E. coli* as a thioredoxin fusion protein with a histidine tag. The trx-P501S fusion protein was purified by nickel chromatography, digested with thrombin to remove the trx fragment and further purified by an acid precipitation procedure followed by reverse phase HPLC.

Mice were immunized with truncated P501S protein. Serum bleeds from mice that potentially contained anti-P501S polyclonal sera were tested for P501S-specific reactivity using ELISA assays with purified P501S and trx-P501S proteins. Serum bleeds that appeared to react specifically with P501S were then screened for P501S reactivity by Western analysis. Mice that contained a P501S-specific antibody component were sacrificed and spleen cells were used to generate anti-P501S antibody producing hybridomas using standard techniques. Hybridoma supernatants were tested for P501S-specific reactivity initially by ELISA, and subsequently by FACS analysis of reactivity with P501S transduced cells. Based on these results, a monoclonal hybridoma referred to as 10E3 was chosen for further subcloning. A number of subclones were generated, tested for specific reactivity to P501S using ELISA and typed for IgG isotype. The results of this

analysis are shown below in Table V. Of the 16 subclones tested, the monoclonal antibody 10E3-G4-D3 was selected for further study.

Table V

5 Isotype analysis of murine anti-P501S monoclonal antibodies

Hybridoma clone	Isotype	Estimated [Ig] in supernatant (µg/ml)
4D11	IgG1	14.6
1G1	IgG1	0.6
4F6	IgG1	72
4H5	IgG1	13.8
4H5-E12	IgG1	10.7
4H5-EH2	IgG1	9.2
4H5-H2-A10	IgG1	10
4H5-H2-A3	IgG1	12.8
4H5-H2-A10-G6	IgG1	13.6
4H5-H2-B11	IgG1	12.3
10E3	IgG2a	3.4
10E3-D4	IgG2a	3.8
10E3-D4-G3	IgG2a	9.5
10E3-D4-G6	IgG2a	10.4
10E3-E7	IgG2a	6.5
8H12	IgG2a	0.6

The specificity of 10E3-G4-D3 for P501S was examined by FACS analysis. Specifically, cells were fixed (2% formaldehyde, 10 minutes), permeabilized (0.1% saponin, 10 minutes) and stained with 10E3-G4-D3 at 0.5 – 1 µg/ml, followed by incubation with a secondary, FITC-conjugated goat anti-mouse Ig antibody (Pharmin-
 10 San Diego, CA). Cells were then analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. For analysis of infected cells, B-LCL were infected
 15 with a vaccinia vector that expresses P501S. To demonstrate specificity in these assays, B-LCL transduced with a different antigen (P703P) and uninfected B-LCL vectors were

utilized. 10E3-G4-D3 was shown to bind with P501S-transduced B-LCL and also with P501S-infected B-LCL, but not with either uninfected cells or P703P-transduced cells.

To determine whether the epitope recognized by 10E3-G4-D3 was found on the surface or in an intracellular compartment of cells, B-LCL were transduced with P501S or HLA-B8 as a control antigen and either fixed and permeabilized as described above or directly stained with 10E3-G4-D3 and analyzed as above. Specific recognition of P501S by 10E3-G4-D3 was found to require permeabilization, suggesting that the epitope recognized by this antibody is intracellular.

The reactivity of 10E3-G4-D3 with the three prostate tumor cell lines Lncap, PC-3 and DU-145, which are known to express high, medium and very low levels of P501S, respectively, was examined by permeabilizing the cells and treating them as described above. Higher reactivity of 10E3-G4-D3 was seen with Lncap than with PC-3, which in turn showed higher reactivity than DU-145. These results are in agreement with the real time PCR and demonstrate that the antibody specifically recognizes P501S in these tumor cell lines and that the epitope recognized in prostate tumor cell lines is also intracellular.

Specificity of 10E3-G4-D3 for P501S was also demonstrated by Western blot analysis. Lysates from the prostate tumor cell lines Lncap, DU-145 and PC-3, from P501S-transiently transfected HEK293 cells, and from non-transfected HEK293 cells were generated. Western blot analysis of these lysates with 10E3-G4-D3 revealed a 46 kDa immunoreactive band in Lncap, PC-3 and P501S-transfected HEK cells, but not in DU-145 cells or non-transfected HEK293 cells. P501S mRNA expression is consistent with these results since semi-quantitative PCR analysis revealed that P501S mRNA is expressed in Lncap, to a lesser but detectable level in PC-3 and not at all in DU-145 cells. Bacterially expressed and purified recombinant P501S (referred to as P501SStr2) was recognized by 10E3-G4-D3 (24 kDa), as was full-length P501S that was transiently expressed in HEK293 cells using either the expression vector VR1012 or pCEP4. Although the predicted molecular weight of P501S is 60.5 kDa, both transfected and “native” P501S run at a slightly lower mobility due to its hydrophobic nature.

Immunohistochemical analysis was performed on prostate tumor and a panel of normal tissue sections (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). Tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before
 5 being sliced into 10 micron sections. Tissue sections were permeabilized and incubated with 10E3-G4-D3 antibody for 1 hr. HRP-labeled anti-mouse followed by incubation with DAB chromogen was used to visualize P501S immunoreactivity. P501S was found to be highly expressed in both normal prostate and prostate tumor tissue but was not detected in any of the other tissues tested.

10 To identify the epitope recognized by 10E3-G4-D3, an epitope mapping approach was pursued. A series of 13 overlapping 20-21 mers (5 amino acid overlap; SEQ ID NO: 489-501) was synthesized that spanned the fragment of P501S used to generate 10E3-G4-D3. Flat bottom 96 well microtiter plates were coated with either the peptides or the P501S fragment used to immunize mice, at 1 microgram/ml for 2 hours at 37 °C. Wells
 15 were then aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified antibody 10E3-G4-D3 was added at 2 fold dilutions (1000 ng – 16 ng) in PBST and incubated for 30 minutes at room temperature. This was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti-mouse
 20 IgG (H+L)Affinipure F(ab') fragment (Jackson ImmunoResearch, West Grove, PA) at 1:20000 for 30 minutes. Plates were then washed and incubated for 15 minutes in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 8, reactivity was seen with the peptide of SEQ ID NO: 496 (corresponding to amino acids 439-459 of
 25 P501S) and with the P501S fragment but not with the remaining peptides, demonstrating that the epitope recognized by 10E3-G4-D3 is localized to amino acids 439-459 of SEQ ID NO: 113.

In order to further evaluate the tissue specificity of P501S, multi-array immunohistochemical analysis was performed on approximately 4700 different human

tissues encompassing all the major normal organs as well as neoplasias derived from these tissues. Sixty-five of these human tissue samples were of prostate origin. Tissue sections 0.6 mm in diameter were formalin-fixed and paraffin embedded. Samples were pretreated with HIER using 10 mM citrate buffer pH 6.0 and boiling for 10 min. Sections were stained with 10E3-G4-D3 and P501S immunoreactivity was visualized with HRP. All the 65 prostate tissues samples (5 normal, 55 untreated prostate tumors, 5 hormone refractory prostate tumors) were positive, showing distinct perinuclear staining. All other tissues examined were negative for P501S expression.

c) Preparation and Characterization of Antibodies against P503S

A fragment of P503S (amino acids 113-241 of SEQ ID NO: 114) was expressed and purified from bacteria essentially as described above for P501S and used to immunize both rabbits and mice. Mouse monoclonal antibodies were isolated using standard hybridoma technology as described above. Rabbit monoclonal antibodies were isolated using Selected Lymphocyte Antibody Method (SLAM) technology at Immgenics Pharmaceuticals (Vancouver, BC, Canada). Table VI, below, lists the monoclonal antibodies that were developed against P503S.

Table VI

Antibody	Species
20D4	Rabbit
JA1	Rabbit
1A4	Mouse
1C3	Mouse
1C9	Mouse
1D12	Mouse
2A11	Mouse
2H9	Mouse
4H7	Mouse
8A8	Mouse
8D10	Mouse
9C12	Mouse

Antibody	Species
6D12	Mouse

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 20D4 and JA1 were determined and are provided in SEQ ID NO: 502 and 503, respectively.

5 In order to better define the epitope binding region of each of the antibodies, a series of overlapping peptides were generated that span amino acids 109-213 of SEQ ID NO: 114. These peptides were used to epitope map the anti-P503S monoclonal antibodies by ELISA as follows. The recombinant fragment of P503S that was employed as the immunogen was used as a positive control. Ninety-six well microtiter plates were coated
10 with either peptide or recombinant antigen at 20 ng/well overnight at 4 °C. Plates were aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature then washed in PBS containing 0.1% Tween 20 (PBST). Purified rabbit monoclonal antibodies diluted in PBST were added to the wells and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and
15 incubation with Protein-A HRP conjugate at a 1:2000 dilution for a further 30 min. Plates were washed six times in PBST and incubated with tetramethylbenzidine (TMB) substrate for a further 15 min. The reaction was stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using at ELISA plate reader. ELISA with the mouse monoclonal antibodies was performed with supernatants from tissue culture run neat in the
20 assay.

All of the antibodies bound to the recombinant P503S fragment, with the exception of the negative control SP2 supernatant. 20D4, JA1 and 1D12 bound strictly to peptide #2101 (SEQ ID NO: 504), which corresponds to amino acids 151-169 of SEQ ID NO: 114. 1C3 bound to peptide #2102 (SEQ ID NO: 505), which corresponds to amino
25 acids 165-184 of SEQ ID NO: 114. 9C12 bound to peptide #2099 (SEQ ID NO: 522), which corresponds to amino acids 120-139 of SEQ ID NO: 114. The other antibodies bind to regions that were not examined in these studies.

Subsequent to epitope mapping, the antibodies were tested by FACS analysis on a cell line that stably expressed P503S to confirm that the antibodies bind to cell surface epitopes. Cells stably transfected with a control plasmid were employed as a negative control. Cells were stained live with no fixative. 0.5 ug of anti-P503S monoclonal antibody was added and cells were incubated on ice for 30 min before being washed twice and incubated with a FITC-labelled goat anti-rabbit or mouse secondary antibody for 20 min. After being washed twice, cells were analyzed with an Excalibur fluorescent activated cell sorter. The monoclonal antibodies 1C3, 1D12, 9C12, 20D4 and JA1, but not 8D3, were found to bind to a cell surface epitope of P503S.

In order to determine which tissues express P503S, immunohistochemical analysis was performed, essentially as described above, on a panel of normal tissues (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). HRP-labeled anti-mouse or anti-rabbit antibody followed by incubation with TMB was used to visualize P503S immunoreactivity. P503S was found to be highly expressed in prostate tissue, with lower levels of expression being observed in cervix, colon, ileum and kidney, and no expression being observed in adrenal, breast, duodenum, gall bladder, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis.

Western blot analysis was used to characterize anti-P503S monoclonal antibody specificity. SDS-PAGE was performed on recombinant (rec) P503S expressed in and purified from bacteria and on lysates from HEK293 cells transfected with full length P503S. Protein was transferred to nitrocellulose and then Western blotted with each of the anti-P503S monoclonal antibodies (20D4, JA1, 1D12, 6D12 and 9C12) at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to either a goat anti-mouse monoclonal antibody or to protein A-sepharose. The monoclonal antibody 20D4 detected the appropriate molecular weight 14 kDa recombinant P503S (amino acids 113-241) and the 23.5 kDa species in the HEK293 cell lysates transfected with full length P503S. Other anti-P503S monoclonal antibodies displayed similar specificity by Western blot.

d) Preparation and Characterization of Antibodies against P703P

Rabbits were immunized with either a truncated (P703Ptr1; SEQ ID NO: 172) or full-length mature form (P703Pfl; SEQ ID NO: 523) of recombinant P703P protein was expressed in and purified from bacteria as described above. Affinity purified polyclonal antibody was generated using immunogen P703Pfl or P703Ptr1 attached to a solid support. Rabbit monoclonal antibodies were isolated using SLAM technology at Immgenics Pharmaceuticals. Table VII below lists both the polyclonal and monoclonal antibodies that were generated against P703P.

Table VII

Antibody	Immunogen	Species/type
Aff. Purif. P703P (truncated); #2594	P703Ptr1	Rabbit polyclonal
Aff. Purif. P703P (full length); #9245	P703Pfl	Rabbit polyclonal
2D4	P703Ptr1	Rabbit monoclonal
8H2	P703Ptr1	Rabbit monoclonal
7H8	P703Ptr1	Rabbit monoclonal

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 8H2, 7H8 and 2D4 were determined and are provided in SEQ ID NO: 506-508, respectively.

Epitope mapping studies were performed as described above. Monoclonal antibodies 2D4 and 7H8 were found to specifically bind to the peptides of SEQ ID NO: 509 (corresponding to amino acids 145-159 of SEQ ID NO: 172) and SEQ ID NO: 510 (corresponding to amino acids 11-25 of SEQ ID NO: 172), respectively. The polyclonal antibody 2594 was found to bind to the peptides of SEQ ID NO: 511-514, with the polyclonal antibody 9427 binding to the peptides of SEQ ID NO: 515-517.

The specificity of the anti-P703P antibodies was determined by Western blot analysis as follows. SDS-PAGE was performed on (1) bacterially expressed recombinant antigen; (2) lysates of HEK293 cells and Ltk^{-/-} cells either untransfected or

transfected with a plasmid expressing full length P703P; and (3) supernatant isolated from these cell cultures. Protein was transferred to nitrocellulose and then Western blotted using the anti-P703P polyclonal antibody #2594 at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to an anti-rabbit antibody. A
 5 35 kDa immunoreactive band could be observed with recombinant P703P. Recombinant P703P runs at a slightly higher molecular weight since it is epitope tagged. In lysates and supernatants from cells transfected with full length P703P, a 30 kDa band corresponding to P703P was observed. To assure specificity, lysates from HEK293 cells stably transfected with a control plasmid were also tested and were negative for P703P expression. Other
 10 anti-P703P antibodies showed similar results.

Immunohistochemical studies were performed as described above, using anti-P703P monoclonal antibody. P703P was found to be expressed at high levels in normal prostate and prostate tumor tissue but was not detectable in all other tissues tested (breast tumor, lung tumor and normal kidney).

15 e) Preparation and Characterization of Antibodies against P504S

Full-length P504S (SEQ ID NO: 108) was expressed and purified from bacteria essentially as described above for P501S and employed to raise rabbit monoclonal antibodies using Selected Lymphocyte Antibody Method (SLAM) technology at Immgenics Pharmaceuticals (Vancouver, BC, Canada). The anti-P504S monoclonal
 20 antibody 13H4 was shown by Western blot to bind to both expressed recombinant P504S and to native P504S in tumor cells.

Immunohistochemical studies using 13H4 to assess P504S expression in various prostate tissues were performed as described above. A total of 104 cases, including 65 cases of radical prostatectomies with prostate cancer (PC), 26 cases of prostate biopsies
 25 and 13 cases of benign prostate hyperplasia (BPH), were stained with the anti-P504S monoclonal antibody 13H4. P504S showed strongly cytoplasmic granular staining in 64/65 (98.5%) of PCs in prostatectomies and 26/26 (100%) of PCs in prostatic biopsies. P504S was stained strongly and diffusely in carcinomas (4+ in 91.2% of cases of PC; 3+ in

5.5%; 2+ in 2.2% and 1+ in 1.1%) and high grade prostatic intraepithelial neoplasia (4+ in all cases). The expression of P504S did not vary with Gleason score. Only 17/91 (18.7%) of cases of NP/BPH around PC and 2/13 (15.4%) of BPH cases were focally (1+, no 2+ to 4+ in all cases) and weakly positive for P504S in large glands. Expression of P504S was not found in small atrophic glands, postatrophic hyperplasia, basal cell hyperplasia and transitional cell metaplasia in either biopsies or prostatectomies. P504S was thus found to be over-expressed in all Gleason scores of prostate cancer (98.5 to 100% of sensitivity) and exhibited only focal positivities in large normal glands in 19/104 of cases (82.3% of specificity). These findings indicate that P504S may be usefully employed for the diagnosis of prostate cancer.

EXAMPLE 19

CHARACTERIZATION OF CELL SURFACE EXPRESSION AND CHROMOSOME LOCALIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes studies demonstrating that the prostate-specific antigen P501S is expressed on the surface of cells, together with studies to determine the probable chromosomal location of P501S.

The protein P501S (SEQ ID NO: 113) is predicted to have 11 transmembrane domains. Based on the discovery that the epitope recognized by the anti-P501S monoclonal antibody 10E3-G4-D3 (described above in Example 17) is intracellular, it was predicted that following transmembrane determinants would allow the prediction of extracellular domains of P501S. Fig. 9 is a schematic representation of the P501S protein showing the predicted location of the transmembrane domains and the intracellular epitope described in Example 17. Underlined sequence represents the predicted transmembrane domains, bold sequence represents the predicted extracellular domains, and italicized sequence represents the predicted intracellular domains. Sequence that is both bold and underlined represents sequence employed to generate polyclonal rabbit serum. The location of the transmembrane domains was predicted using HHMTOP as described by

Tusnady and Simon (Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to Topology Prediction, *J. Mol. Biol.* 283:489-506, 1998).

Based on Fig. 9, the P501S domain flanked by the transmembrane domains
 5 corresponding to amino acids 274-295 and 323-342 is predicted to be extracellular. The
 peptide of SEQ ID NO: 518 corresponds to amino acids 306-320 of P501S and lies in the
 predicted extracellular domain. The peptide of SEQ ID NO: 519, which is identical to the
 peptide of SEQ ID NO: 518 with the exception of the substitution of the histidine with an
 asparagine, was synthesized as described above. A Cys-Gly was added to the C-terminus
 10 of the peptide to facilitate conjugation to the carrier protein. Cleavage of the peptide from
 the solid support was carried out using the following cleavage mixture: trifluoroacetic
 acid:ethanediol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for two hours, the
 peptide was precipitated in cold ether. The peptide pellet was then dissolved in 10% v/v
 acetic acid and lyophilized prior to purification by C18 reverse phase hplc. A gradient of
 15 5-60% acetonitrile (containing 0.05% TFA) in water (containing 0.05% TFA) was used to
 elute the peptide. The purity of the peptide was verified by hplc and mass spectrometry,
 and was determined to be >95%. The purified peptide was used to generate rabbit
 polyclonal antisera as described above.

Surface expression of P501S was examined by FACS analysis. Cells were
 20 stained with the polyclonal anti-P501S peptide serum at 10 µg/ml, washed, incubated with
 a secondary FITC-conjugated goat anti-rabbit Ig antibody (ICN), washed and analyzed for
 FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS
 analysis of transduced cells, B-LCL were retrovirally transduced with P501S. To
 demonstrate specificity in these assays, B-LCL transduced with an irrelevant antigen
 25 (P703P) or nontransduced were stained in parallel. For FACS analysis of prostate tumor
 cell lines, Lncap, PC-3 and DU-145 were utilized. Prostate tumor cell lines were
 dissociated from tissue culture plates using cell dissociation medium and stained as above.
 All samples were treated with propidium iodide (PI) prior to FACS analysis, and data was
 obtained from PI-excluding (*i.e.*, intact and non-permeabilized) cells. The rabbit

polyclonal serum generated against the peptide of SEQ ID NO: 519 was shown to specifically recognize the surface of cells transduced to express P501S, demonstrating that the epitope recognized by the polyclonal serum is extracellular.

To determine biochemically if P501S is expressed on the cell surface, peripheral membranes from Lncap cells were isolated and subjected to Western blot analysis. Specifically, Lncap cells were lysed using a dounce homogenizer in 5 ml of homogenization buffer (250 mM sucrose, 10 mM HEPES, 1mM EDTA, pH 8.0, 1 complete protease inhibitor tablet (Boehringer Mannheim)). Lysate samples were spun at 1000 g for 5 min at 4 °C. The supernatant was then spun at 8000g for 10 min at 4 °C. Supernatant from the 8000g spin was recovered and subjected to a 100,000g spin for 30 min at 4 °C to recover peripheral membrane. Samples were then separated by SDS-PAGE and Western blotted with the mouse monoclonal antibody 10E3-G4-D3 (described above in Example 17) using conditions described above. Recombinant purified P501S, as well as HEK293 cells transfected with and over-expressing P501S were included as positive controls for P501S detection. LCL cell lysate was included as a negative control. P501S could be detected in Lncap total cell lysate, the 8000g (internal membrane) fraction and also in the 100,000g (plasma membrane) fraction. These results indicate that P501S is expressed at, and localizes to, the peripheral membrane.

To demonstrate that the rabbit polyclonal antiserum generated to the peptide of SEQ ID NO: 519 specifically recognizes this peptide as well as the corresponding native peptide of SEQ ID NO: 518, ELISA analyses were performed. For these analyses, flat-bottomed 96 well microtiter plates were coated with either the peptide of SEQ ID NO: 519, the longer peptide of SEQ ID NO: 520 that spans the entire predicted extracellular domain, the peptide of SEQ ID NO: 521 which represents the epitope recognized by the P501S-specific antibody 10E3-G4-D3, or a P501S fragment (corresponding to amino acids 355-526 of SEQ ID NO: 113) that does not include the immunizing peptide sequence, at 1 µg/ml for 2 hours at 37 °C. Wells were aspirated, blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified anti-P501S polyclonal rabbit serum was

added at 2 fold dilutions (1000 ng - 125 ng) in PBST and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubating with HRP-conjugated goat anti-rabbit IgG (H+L) Affinipure F(ab') fragment at 1:20000 for 30 min. Plates were then washed and incubated for 15 min in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 11, the anti-P501S polyclonal rabbit serum specifically recognized the peptide of SEQ ID NO: 519 used in the immunization as well as the longer peptide of SEQ ID NO: 520, but did not recognize the irrelevant P501S-derived peptides and fragments.

In further studies, rabbits were immunized with peptides derived from the P501S sequence and predicted to be either extracellular or intracellular, as shown in Fig. 9. Polyclonal rabbit sera were isolated and polyclonal antibodies in the serum were purified, as described above. To determine specific reactivity with P501S, FACS analysis was employed, utilizing either B-LCL transduced with P501S or the irrelevant antigen P703P, of B-LCL infected with vaccinia virus-expressing P501S. For surface expression, dead and non-intact cells were excluded from the analysis as described above. For intracellular staining, cells were fixed and permeabilized as described above. Rabbit polyclonal serum generated against the peptide of SEQ ID NO: 548, which corresponds to amino acids 181-198 of P501S, was found to recognize a surface epitope of P501S. Rabbit polyclonal serum generated against the peptide SEQ ID NO: 551, which corresponds to amino acids 543-553 of P501S, was found to recognize an epitope that was either potentially extracellular or intracellular since in different experiments intact or permeabilized cells were recognized by the polyclonal sera. Based on similar deductive reasoning, the sequences of SEQ ID NO: 541-547, 549 and 550, which correspond to amino acids 109-122, 539-553, 509-520, 37-54, 342-359, 295-323, 217-274, 143-160 and 75-88, respectively, of P501S, can be considered to be potential surface epitopes of P501S recognized by antibodies.

The chromosomal location of P501S was determined using the GeneBridge 4 Radiation Hybrid panel (Research Genetics). The PCR primers of SEQ ID NO: 528 and

529 were employed in PCR with DNA pools from the hybrid panel according to the manufacturer's directions. After 38 cycles of amplification, the reaction products were separated on a 1.2% agarose gel, and the results were analyzed through the Whitehead Institute/MIT Center for Genome Research web server (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) to determine the probable chromosomal location. Using this approach, P501S was mapped to the long arm of chromosome 1 at WI-9641 between q32 and q42. This region of chromosome 1 has been linked to prostate cancer susceptibility in hereditary prostate cancer (Smith *et al. Science* 274:1371-1374, 1996 and Berthon *et al. Am. J. Hum. Genet.* 62:1416-1424, 1998). These results suggest that P501S may play a role in prostate cancer malignancy.

EXAMPLE 20

REGULATION OF EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

15 Steroid (androgen) hormone modulation is a common treatment modality in prostate cancer. The expression of a number of prostate tissue-specific antigens have previously been demonstrated to respond to androgen. The responsiveness of the prostate-specific antigen P501S to androgen treatment was examined in a tissue culture system as follows.

20 Cells from the prostate tumor cell line LNCaP were plated at 1.5×10^6 cells/T75 flask (for RNA isolation) or 3×10^5 cells/well of a 6-well plate (for FACS analysis) and grown overnight in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum (BRL Life Technologies, Gaithersburg, MD). Cell culture was continued for an additional 72 hours in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum, with 1 nM of the synthetic androgen Methyltrienolone (R1881; New England Nuclear) added at various time points. Cells were then harvested for RNA isolation and FACS analysis at 0, 1, 2, 4, 8, 16, 24, 28 and 72-hours post androgen addition. FACS analysis was performed using the anti-P501S antibody 10E3-G4-D3 and permeabilized cells.

For Northern analysis, 5-10 micrograms of total RNA was run on a formaldehyde denaturing gel, transferred to Hybond-N nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ), cross-linked and stained with methylene blue. The filter was then prehybridized with Church's Buffer (250 mM Na₂HPO₄, 70 mM H₃PO₄, 1 mM EDTA, 1% SDS, 1% BSA in pH 7.2) at 65 °C for 1 hour. P501S DNA was labeled with ³²P using High Prime random-primed DNA labeling kit (Boehringer Mannheim). Unincorporated label was removed using MicroSpin S300-HR columns (Amersham Pharmacia Biotech). The RNA filter was then hybridized with fresh Church's Buffer containing labeled cDNA overnight, washed with 1X SCP (0.1 M NaCl, 0.03 M Na₂HPO₄·7H₂O, 0.001 M Na₂EDTA), 1% sarkosyl (n-lauroylsarcosine) and exposed to X-ray film.

Using both FACS and Northern analysis, P501S message and protein levels were found to increase in response to androgen treatment.

15

EXAMPLE 20

PREPARATION OF FUSION PROTEINS OF PROSTATE-SPECIFIC ANTIGENS

The example describes the preparation of a fusion protein of the prostate-specific antigen P703P and a truncated form of the known prostate antigen PSA. The truncated form of PSA has a 21 amino acid deletion around the active serine site. The expression construct for the fusion protein also has a restriction site at 3' end, immediately prior to the termination codon, to aid in adding cDNA for additional antigens.

The full-length cDNA for PSA was obtained by RT-PCR from a pool of RNA from human prostate tumor tissues using the primers of SEQ ID NO: 607 and 608, and cloned in the vector pCR-Blunt II-TOPO. The resulting cDNA was employed as a template to make two different fragments of PSA by PCR with two sets of primers (SEQ ID NO: 609 and 610; and SEQ ID NO: 611 and 612). The PCR products having the expected size were used as templates to make truncated forms of PSA by PCR with the primers of SEQ ID NO: 611 and 613, which generated PSA (delta 208-218 in amino acids).

The cDNA for the mature form of P703P with a 6X histidine tag at the 5' end, was prepared by PCR with P703P and the primers of SEQ ID NO: 614 and 615. The cDNA for the fusion of P703P with the truncated form of PSA (referred to as FOPP) was then obtained by PCR using the modified P703P cDNA and the truncated form of PSA cDNA as
 5 templates and the primers of SEQ ID NO: 614 and 615. The FOPP cDNA was cloned into the NdeI site and XhoI site of the expression vector pCRX1, and confirmed by DNA sequencing. The determined cDNA sequence for the fusion construct FOPP is provided in SEQ ID NO: 616, with the amino acid sequence being provided in SEQ ID NO: 617.

The fusion FOPP was expressed as a single recombinant protein in *E. coli* as
 10 follows. The expression plasmid pCRX1FOPP was transformed into the *E. coli* strain BL21-CodonPlus RIL. The transformant was shown to express FOPP protein upon induction with 1 mM IPTG. The culture of the corresponding expression clone was inoculated into 25 ml LB broth containing 50 ug/ml kanamycin and 34 ug/ml chloramphenicol, grown at 37 °C to OD600 of about 1, and stored at 4 °C overnight. The
 15 culture was diluted into 1 liter of TB LB containing 50 ug/ml kanamycin and 34 ug/ml chloramphenicol, and grown at 37 °C to OD600 of 0.4. IPTG was added to a final concentration of 1 mM, and the culture was incubated at 30 °C for 3 hours. The cells were pelleted by centrifugation at 5,000 RPM for 8 min. To purify the protein, the cell pellet was suspended in 25 ml of 10 mM Tris-Cl pH 8.0, 2mM PMSF, complete protease
 20 inhibitor and 15 ug lysozyme. The cells were lysed at 4 °C for 30 minutes, sonicated several times and the lysate centrifuged for 30 minutes at 10,000 x g. The precipitate, which contained the inclusion body, was washed twice with 10 mM Tris-Cl pH 8.0 and 1% CHAPS. The inclusion body was dissolved in 40 ml of 10 mM Tris-Cl pH 8.0, 100 mM sodium phosphate and 8 M urea. The solution was bound to 8 ml Ni-NTA (Qiagen) for
 25 one hour at room temperature. The mixture was poured into a 25 ml column and washed with 50 ml of 10 mM Tris-Cl pH 6.3, 100 mM sodium phosphate, 0.5% DOC and 8M urea. The bound protein was eluted with 350 mM imidazole, 10 mM Tris-Cl pH 8.0, 100 mM sodium phosphate and 8 M urea. The fractions containing FOPP proteins were combined and dialyzed extensively against 10 mM Tris-Cl pH 4.6, aliquoted and stored at - 70 °C.

EXAMPLE 21

REAL-TIME PCR CHARACTERIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S IN
PERIPHERAL BLOOD OF PROSTATE CANCER PATIENTS

5

Circulating epithelial cells were isolated from fresh blood of normal individuals and metastatic prostate cancer patients, mRNA isolated and cDNA prepared using real-time PCR procedures. Real-time PCR was performed with the Taqman™ procedure using both gene specific primers and probes to determine the levels of gene expression.

10

Epithelial cells were enriched from blood samples using an immunomagnetic bead separation method (Dynal A.S., Oslo, Norway). Isolated cells were lysed and the magnetic beads removed. The lysate was then processed for poly A+ mRNA isolation using magnetic beads coated with Oligo(dT)25. After washing the beads in buffer, bead/poly A+ RNA samples were suspended in 10 mM Tris HCl pH 8.0 and subjected to reversed transcription. The resulting cDNA was subjected to real-time PCR using gene specific primers. Beta-actin content was also determined and used for normalization. Samples with P501S copies greater than the mean of the normal samples + 3 standard deviations were considered positive. Real time PCR on blood samples was performed using the Taqman™ procedure but extending to 50 cycles using forward and reverse primers and probes specific for P501S. Of the eight samples tested, 6 were positive for P501S and β -actin signal. The remaining 2 samples had no detectable β -actin or P501S. No P501S signal was observed in the four normal blood samples tested.

15

20

25

EXAMPLE 22

EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGENS P703P AND P501S IN SCID MOUSE-
PASSAGED PROSTATE TUMORS

When considering the effectiveness of antigens in the treatment of prostate cancer, the continued presence of the antigens in tumors during androgen ablation therapy is important. The presence of the prostate-specific antigens P703P and P501S in prostate tumor samples grown in SCID mice in the presence of testosterone was evaluated as follows.

Two prostate tumors that had metastasized to the bone were removed from patients, implanted into SCID mice and grown in the presence of testosterone. Tumors were evaluated for mRNA expression of P703P, P501S and PSA using quantitative real time PCR with the SYBR green assay method. Expression of P703P and P501S in a prostate tumor was used as a positive control and the absence in normal intestine and normal heart as negative controls. In both cases, the specific mRNA was present in late passage tumors. Since the bone metastases were grown in the presence of testosterone, this implies that the presence of these genes would not be lost during androgen ablation therapy.

EXAMPLE 23

ANTI-P503S MONOCLONAL ANTIBODY INHIBITS TUMOR GROWTH *IN VIVO*

The ability of the anti-P503S monoclonal antibody 20D4 to suppress tumor formation in mice was examined as follows.

Ten SCID mice were injected subcutaneously with HEK293 cells that expressed P503S. Five mice received 150 micrograms of 20D4 intravenously at day 0 (time of tumor cell injection), day 5 and day 9. Tumor size was measured for 50 days. Of the five animals that received no 20D4, three formed detectable tumors after about 2 weeks which continued to enlarge throughout the study. In contrast, none of the five mice that received 20D4 formed tumors. These results demonstrate that the anti-P503S Mab 20D4 displays potent anti-tumor activity *in vivo*.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration,

195

CLAIMS

What is Claimed:

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(b) complements of the sequences provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(d) sequences that hybridize to a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(f) sequences having at least 90% identity to a sequence of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381,

382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824; and

(g) degenerate variants of a sequence provided in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877;

(b) sequences having at least 70% identity to a sequence of SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877;

(c) sequences having at least 90% identity to a sequence of SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826, 827, 853, 855, 858, 860-862 and 866-877;

(d) sequences encoded by a polynucleotide of claim 1;

(e) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

(f) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector

5. An isolated antibody, or antigen-binding fragment thereof, that

6. A method for detecting the presence of a cancer in a patient,

(a) obtaining a biological sample from the patient;

(b) contacting the biological sample with a binding agent that binds to a

(c) detecting in the sample an amount of polypeptide that binds to the

(d) comparing the amount of polypeptide to a predetermined cut-off

7. A fusion protein comprising at least one polypeptide according to

8. An oligonucleotide that hybridizes to a sequence recited in SEQ ID

9. A method for stimulating and/or expanding T cells specific for a

(a) polypeptides according to claim 2;

(b) polynucleotides according to claim 1; and

(c) antigen-presenting cells that express a polypeptide according to claim 1,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.

11. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) polynucleotides according to claim 1;
- (c) antibodies according to claim 5;
- (d) fusion proteins according to claim 7;
- (e) T cell populations according to claim 10; and
- (f) antigen presenting cells that express a polypeptide according to claim 2.

12. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 11.

13. A method for the treatment of a cancer in a patient, comprising administering to the patient a composition of claim 11.

14. A method for determining the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;

(b) contacting the biological sample with an oligonucleotide according to claim 8;

(c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

15. A diagnostic kit comprising at least one oligonucleotide according to claim 8.

16. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

17. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;

(b) administering to the patient an effective amount of the proliferated T cells,

and thereby inhibiting the development of a cancer in the patient.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF
PROSTATE CANCER

ABSTRACT OF THE DISCLOSURE

Compositions and methods for the therapy and diagnosis of cancer, particularly prostate cancer, are disclosed. Illustrative compositions comprise one or more prostate-specific polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly prostate cancer.

003050" 6245950

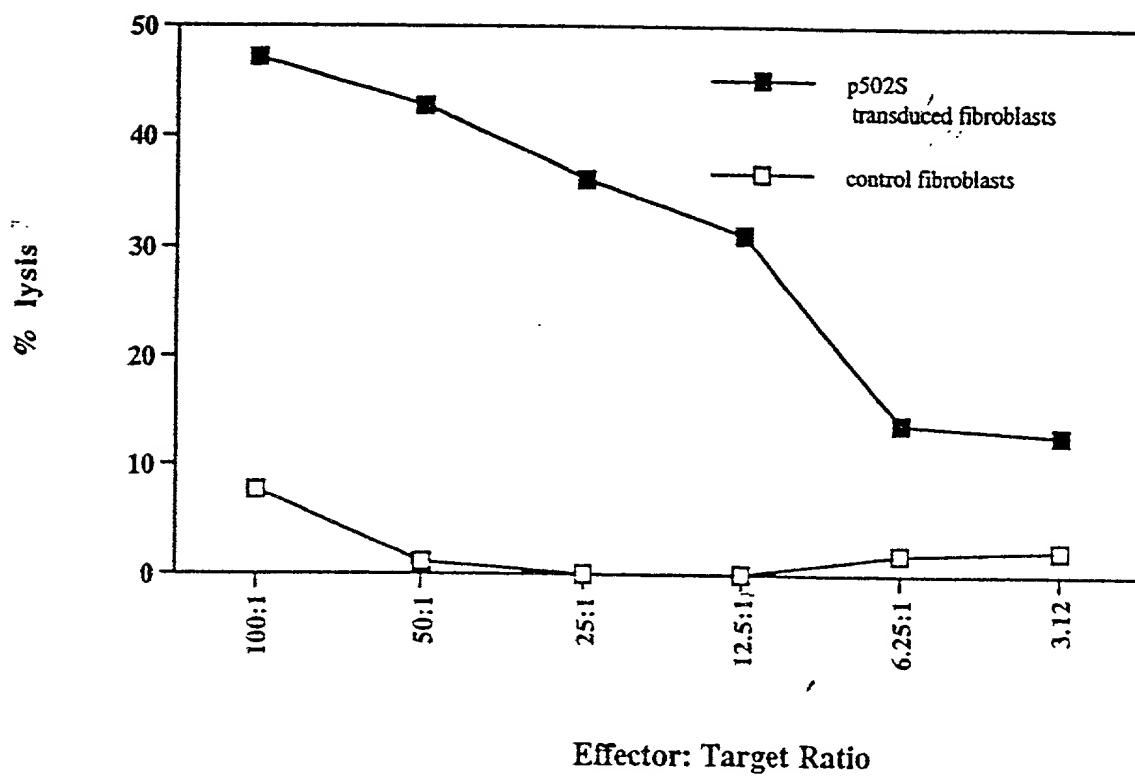


FIG. 1

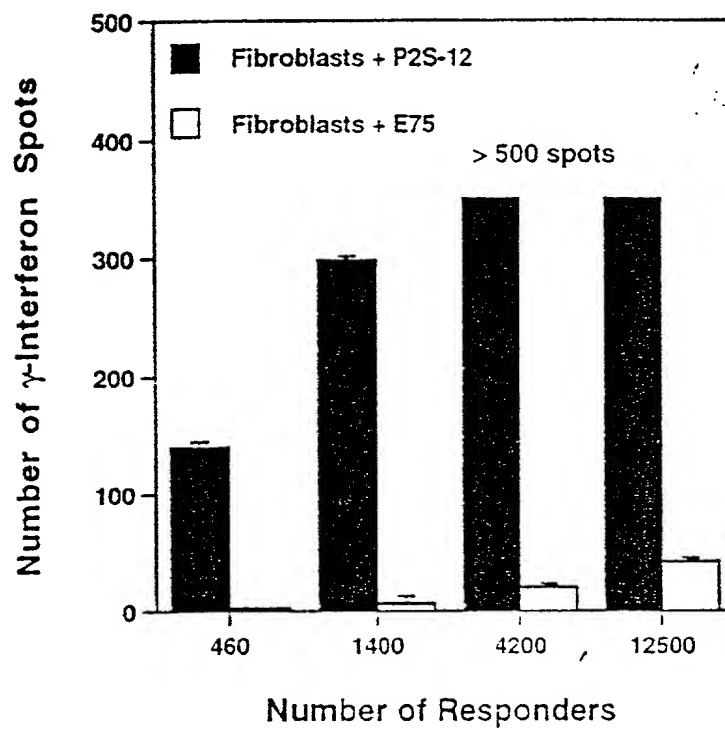


FIG. 2A

00000" 6225000

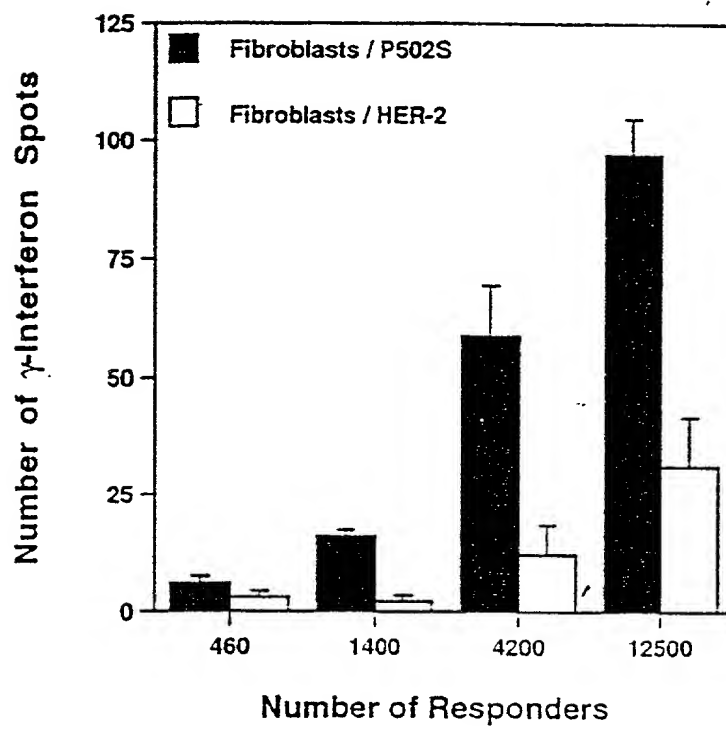
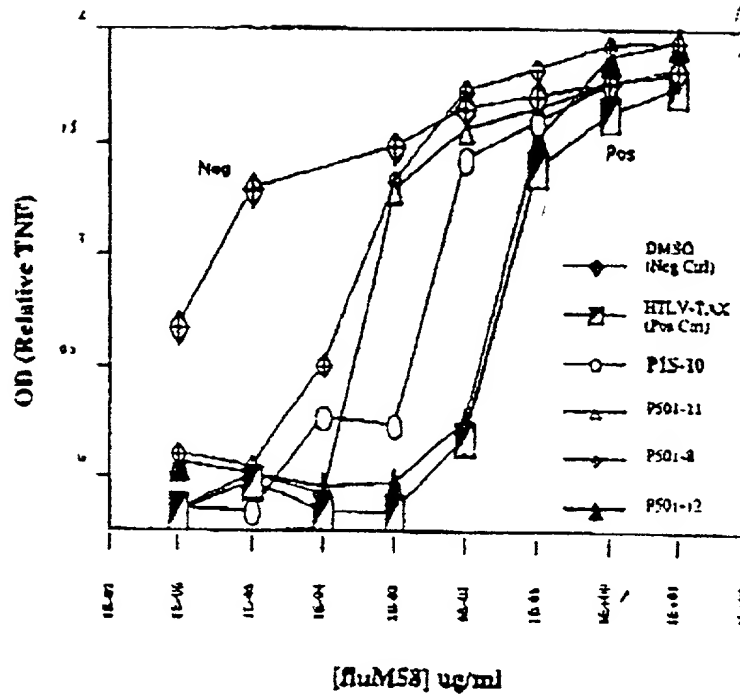


FIG. 2B

009060" 6225960



Figure

3

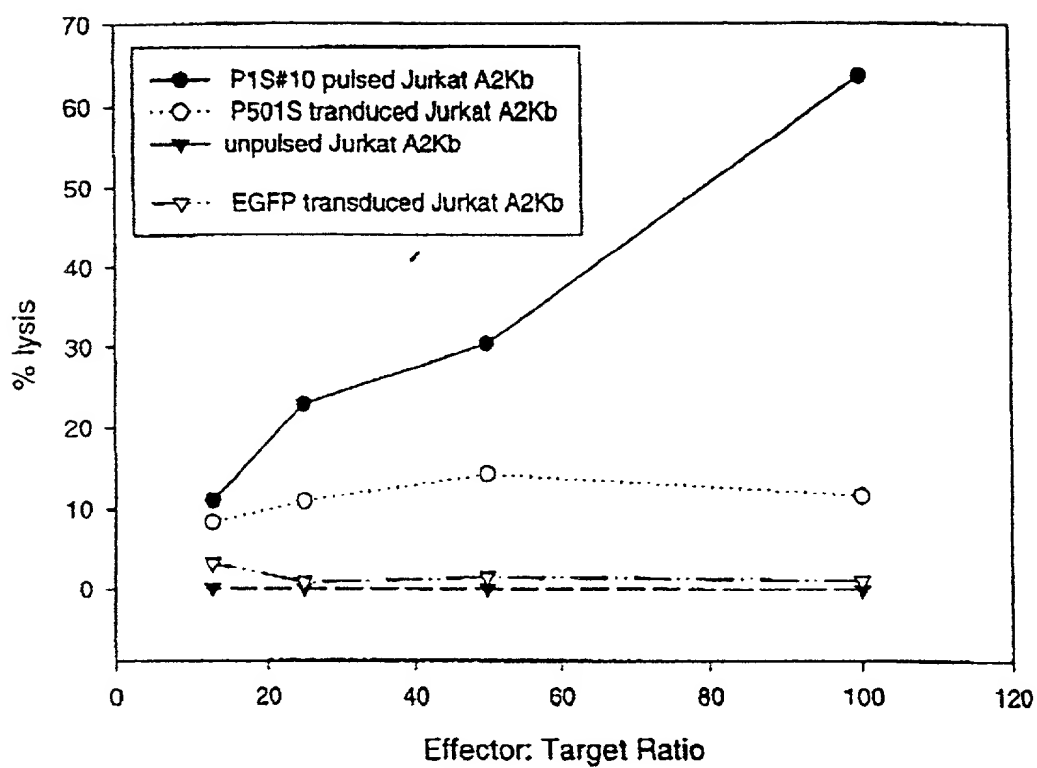


Figure 4

009050" 6225960

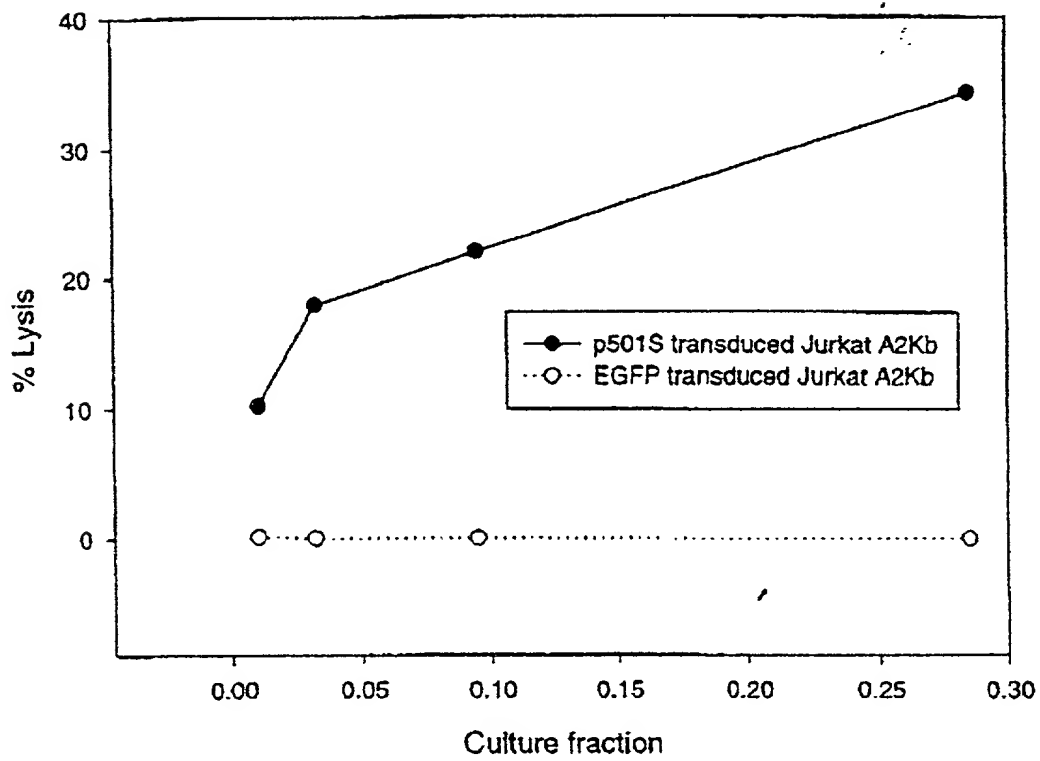


Figure 5

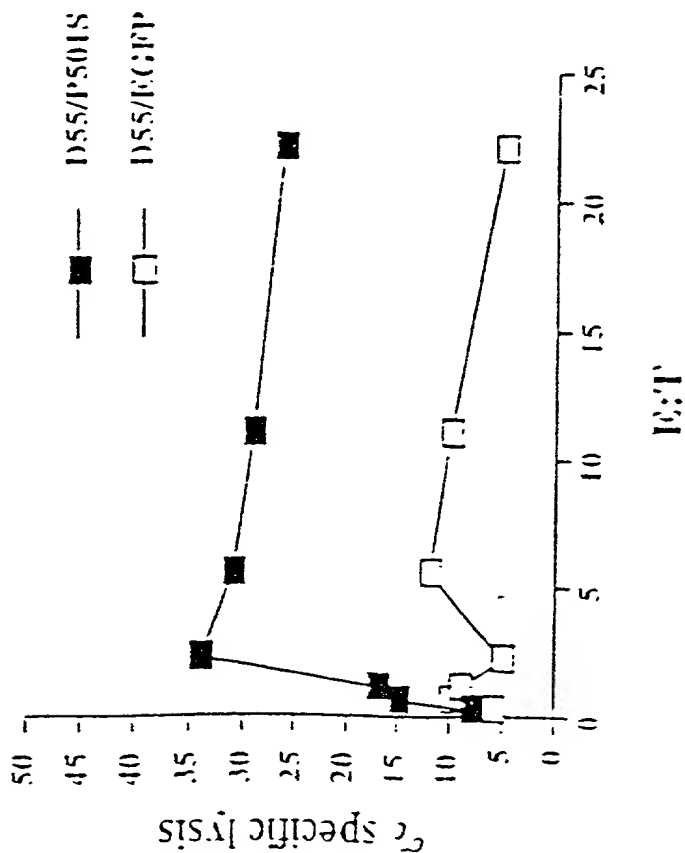


Fig. 6A

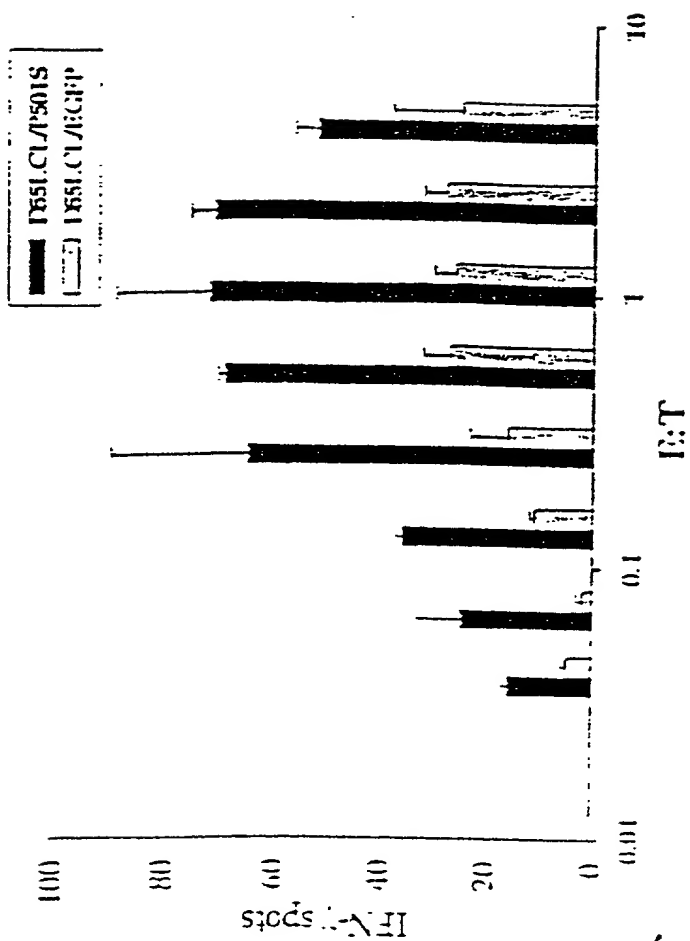
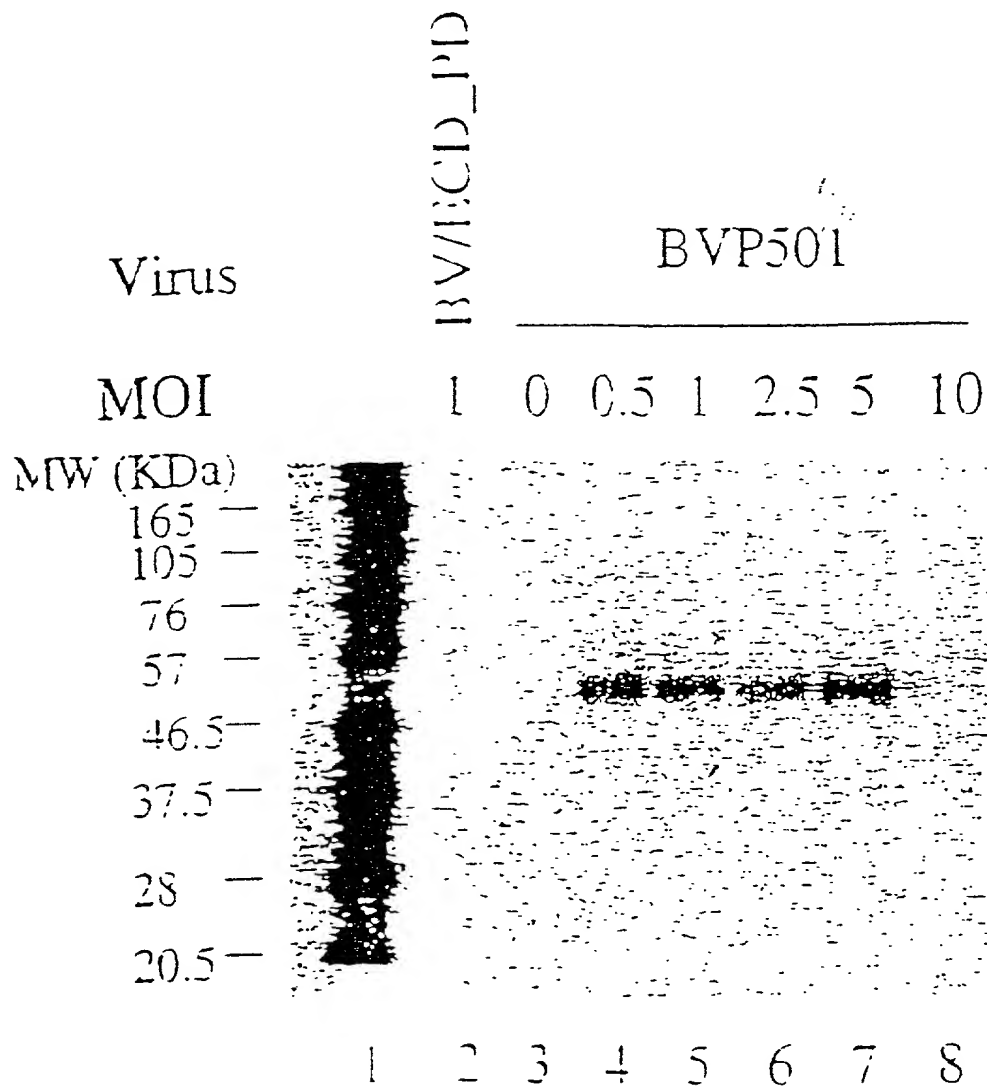


Fig. 6B

Expression of P501S by the Baculovirus Expression System



0.6 million high cells per 6-well plate were infected with an unrelated control virus BV/ECD_PD (lane 2) or without virus (lane 3), or with recombinant baculovirus for P501 at different MOIs (lane 4-8). Cell lysates were run on SDS-PAGE under the reducing condition and analyzed by Western blot with a monoclonal antibody against P501S-ICED-G4D3. Lane 1 is the biotinylated protein molecular weight marker (8 kDa).

Fig. 7

Figure 8. Mapping of the epitope recognized by 10E3-G4-D3

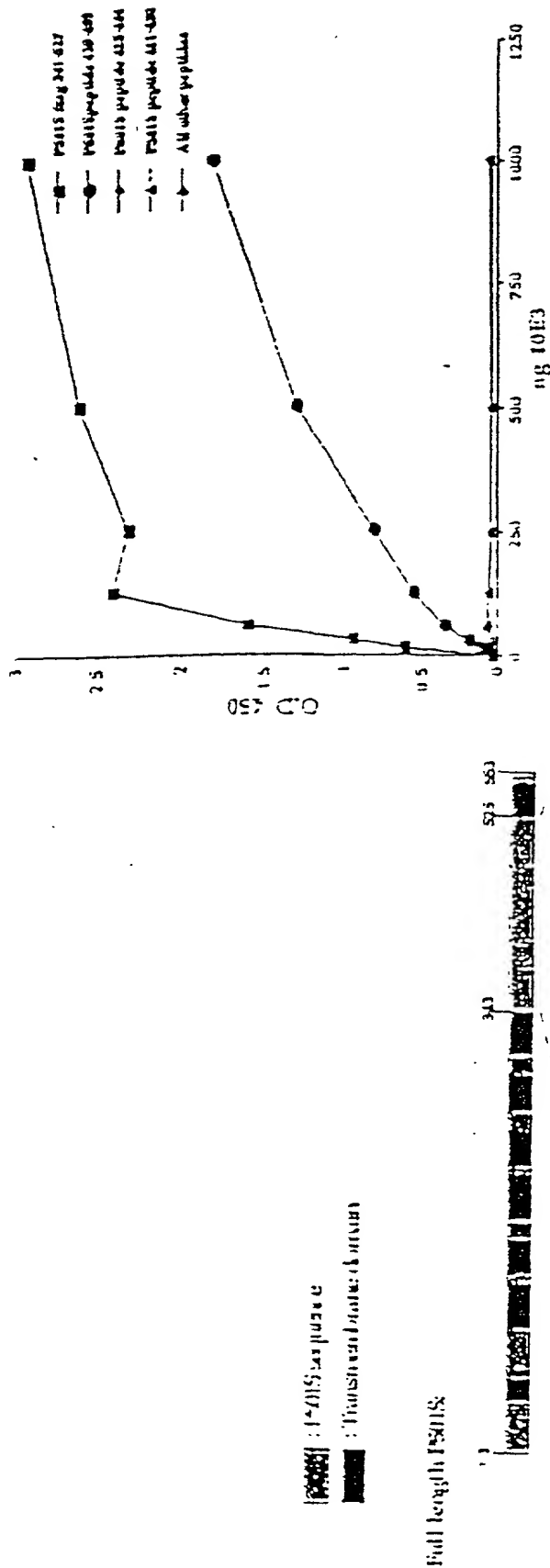
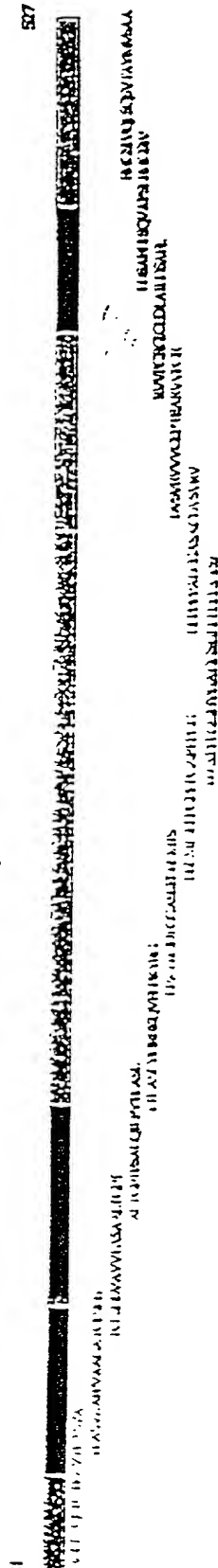


Fig. 8

PSIIS fragment used for immunization:



7

Figure 1. Schematic of P501S with predicted transmembrane, cytoplasmic, and extracellular regions

AVQRLAVSRLLRIIRK AQLILYNLLTTEGLEVCLAAQITVVPPLILLEVGVVEKKFM TMVLGIGPVVLGLVCYPLIGSAS
 DWWRGRYGRRRP FIWALSIGILLSEFLIPRAGWL AGILCTDPRPLE LALLILGVCLLDFCGQVCFITPL
 FALISDLFRDPDHCRQ AYSVYAFHSLGGCTGVITPAI DWVDSATAPVLCIQRE
 CLTGLLTLHLFCVAAITLY AFFVALGPTPEAGLSAPSLSPDCTPCRARAFRNIGALLPRI
 HLLCRAMPRTLRH LPYAFILCSWMAIMETITFYIDP VGEGLYQGGVPRARRGTLEARHHYDEGVR
 MGSILGLFLOCAISLYFSLYM DRAYQRECTRAVYLAS VAAFTYAAGATCLSHSYAVVTA SAA
 LTGELTFSALQILPYTLASLY HREKQVFLPKYRGDTGGASSEDSTMTSEFLPGPKPGAPFPNGHIVGAGCSGL
 JPPPPALCGASACDVSVRVVVGEFTPEARVVVGRG ILLDLAILDSAPILLSQVAPSEF MGSIVQLSQS
 VTAYMVSAAGLGLVAYFAT QVVFDKSLAKYSA

Underlined sequence: Predicted transmembrane domain; Bold sequence: Predicted extracellular domain;
 Italic sequence: Predicted intracellular domain. Sequence in bold/underlined, used to generate polyclonal rabbit serum

Localization of domains predicted using IMMTOPI (G.E. Tusnady and I. Simon (1998) Principles
 Governing Amino Acid Composition of Integral Membrane Proteins: Applications to topology Prediction. J. Mol. Biol. 283,
 489-506.

Genomic Map of (5) Corixa Candidate Genes

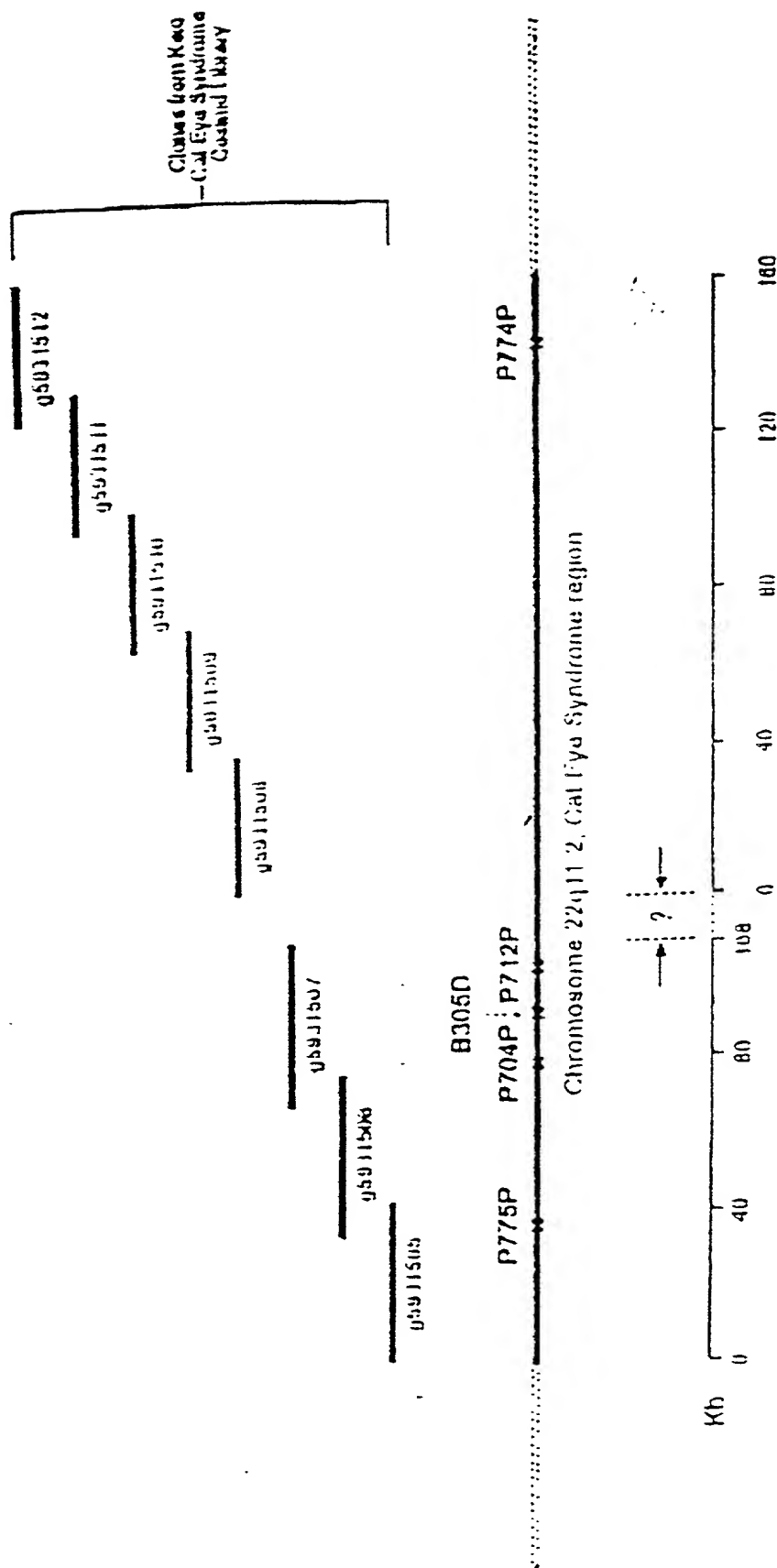


Fig. 10

FIGURE 4. Elisa assay of rabbit polyclonal antibody specificity

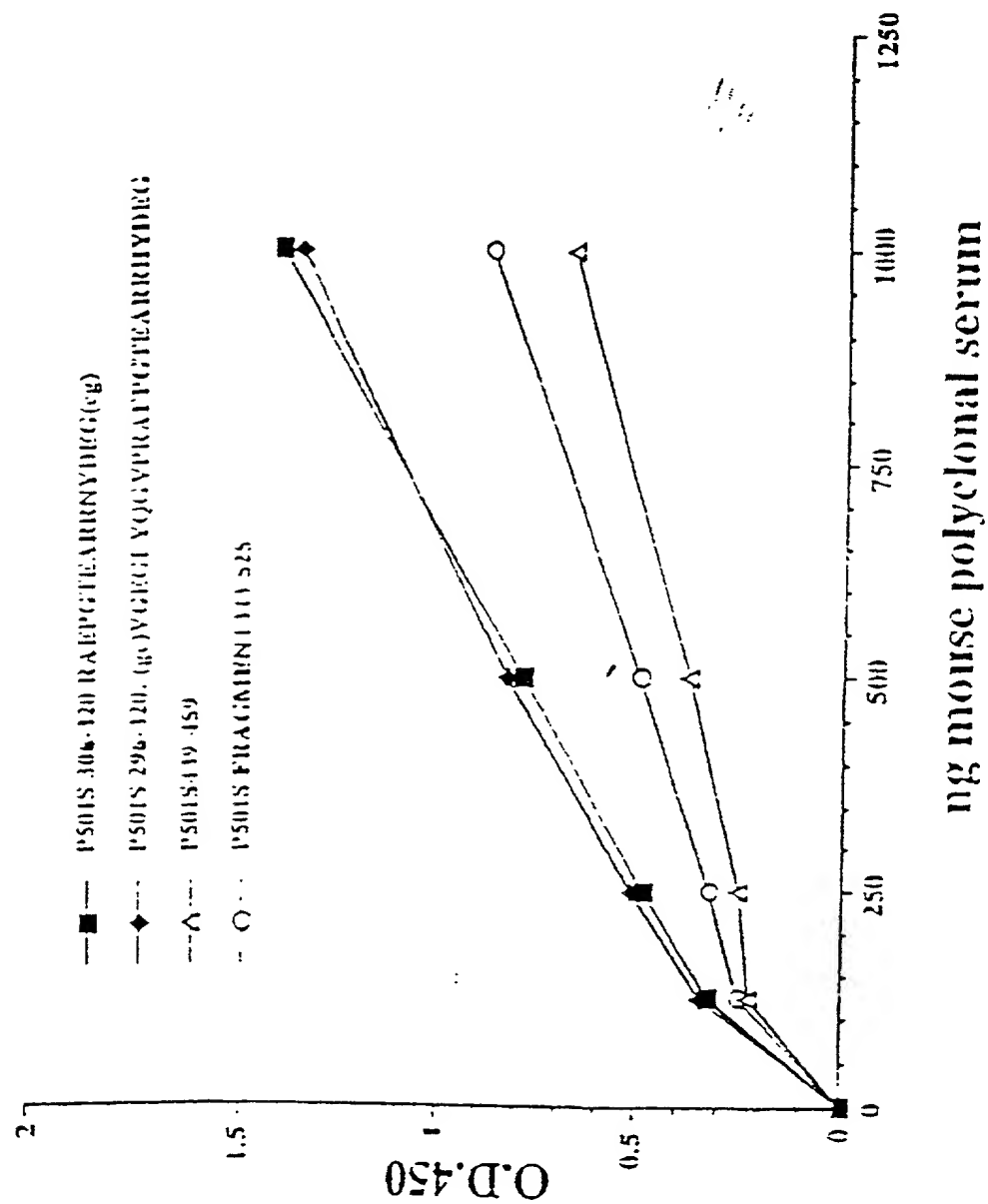


Fig. 11

10	20	30	40	50	60	70
GTCACCTAGGAAAAGGTGTCCTTTCGGGCAGCCGGGCTCAGCATGAGGAACAGAAGGAATGACACTCTGG 70						
ACAGCACCCGGACCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTACACTGAAAGCGACTTGGT 140						
GAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTCTTTACCAAAGATTCCAAGGCCACGGAG 210						
AATGTGTGCAAGTGTGGCTATGCCAGAGCCAGCACATGGAAGGCACCCAGATCAACCAAAGTGAGAAAT 280						
GGAACACAAGAAACACACCAAGGAAATTCCTACCCAGCGCTTTGGGGATATTCAGTTTGAGACACTGGG 350						
360	370	380	390	400	410	420
GAAGAAAGGGAAGTATATACGTCTGTCTTGCACACGACGCGGAAATCCTTTACGAGCTGCTGACCCAG 420						
CACTGGCACCCTGAAAACAACCAACCTGGTCAATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAAGC 490						
CGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGG 560						
AGGCACCCATTATGGCCTGACCAAGTACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCA 630						
GAGGAGAATATTGTGGCCATTGGCATAGCAGCTGGGGCATGGTCTCAACCGGGACACCCCTCATCAGGA 700						
710	720	730	740	750	760	770
ATTGCGATGCTGAGGGCTATTTTTTAGCCCAAGTACCTTATGGATGACTTCACAAGGGATCCACTGTATAT 770						
CCTGGACAACAACCAACACACATTTGCTGCTGGTGGACAAATGGCTGTGATGGACATCCCACTGTGCAAGCA 840						
AAGCTCCGGAATCAGCTAGAGAAAGCATATCTGTGAGCGCACTATTCAAGATTCCAACTATGGTGGCAAGA 910						
TCCCCATTGTGTGTTTGGCCAAAGGAGGTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCAAAAA 980						
TAAAAATTCCTTGTTGTTGGTGGTGGAAAGGCTCGGGCGGGATCGCTGATGTGATCGCTAGCCTGGTGGAGGT 1050						
1060	1070	1080	1090	1100	1110	1120
GAGGATGCCCGGACATCTTCTGCCGTCAAGGAGAAGCTGGTGGCTTTTTTACCCCGCACGGTGTCTCGGC 1120						
TGCTTGAGGAGGAGACTGAGAGTTGGATCAAAATGGCTCAAAAGAAATCTCTCAATGTTCTCACCTATTAA 1190						
AGTTATTAAGATGGAAGAAGCTGGGGATGAAATTTGAGCAATGCCATCTCTACGCTCTATACAAAGCC 1260						
TTCAGCACCAAGTGAGCAAGACAGGATAACTGGAAATGGGC-GCTGAAGCTTCTGCTGGAGTGGAAACAG 1330						
TGGACTTAGCCAAATGATGAGATTTTACCAATGACCGCCGATGGGAGTCTGCTGACCTCAAGAAATCAT 1400						
1410	1420	1430	1440	1450	1460	1470
GTTTACGGCTCTCATAAAGGATAGACCCAAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGG 1470						
AAGTTTCTCACCCATGATGTCTCACTGAAGCTCTCTCCAAGCACTTCAGCACGCTTGTGTACCGGAATC 1540						
TGCAGATCGCCAAGAATTCTATAATGATGCCCTCTCTACGCTTTGTCTGGAAGCTGGTTGCCGAACCTCCG 1610						
AAGAGGCTTCGGGAAGGAAGACAGAAATGGCGGGATGAGATGGACATAGAAGCTCCACGACGTGTCTCT 1680						
ATTACTCGGCACCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAAATAGAAGGAAGCTCTCCAAAG 1750						
1760	1770	1780	1790	1800	1810	1820
TCATTTGGGAGCAGACCCAGGGGCTGCACTCTGSCAGCCCTCGGAGCCAGCAAGCTTCTGAAGACTCTGGC 1820						
CAAAGTGAAGAAACACATCAATGCTGCTGGGGAGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCT 1890						
GTTGAGCTGTCACTGAGTGTACAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTATTCTCTGTG 1960						
AAGCTTGGGGTGAAGCAAGCTGTCTGGAGCTGGCGGTGAAGCCAGAGACCATTTACCGGCCCAAGCC 2030						
TGGGGTCCAGAAATTTCTTTCTAAGCAATGGTATGGAGAGATTTCCGAGACACCAAGAACTGGAAGATT 2100						

Fig. 12A (1)

2110	2120	2130	2140	2150	2160	2170
TCCTGTGTCTGTTTATTATACCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTGACA	2170					
AGCACAAGAAGCTGCTTTGGTACTATGTGGGCTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGT	2240					
GGTCTTCTACATCGCCTTCTCCTGCTGTTTGGCTACGTGCTGCTCATGGATTTCCATTGGTGCCACAC	2310					
CCCCCGAGCTGCTCCTGTACTCCTGGTCTTGTCTCTTCTGTGATGAAGTCAGACAGTGGTACGTAA	2380					
ATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACTTCATAGCAGG	2450					
2460	2470	2480	2490	2500	2510	2520
AATTGTATTTGGGCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACGAGTCATTTTCTGTCTGGAC	2520					
TACATTATTTTCACTCTAAGATTGATCCACATTTTACTGTAAGCAGAAACTTAGGACCCAAGATTATAA	2590					
TGCTGCAGAGGAIGCTGATGATGTGTCTCTTCTGCTCTCTTTGCGGTGTGGATGGTGGCCTTTGG	2660					
CGTGGCCAGGCAAGGGATCCTTAGGCAGAAATGAGCAGCGCTGGAGGTGGATATTCGTTGGTTCATCTAC	2730					
GAGCCCTACCTGGCCATGTTTGGGCCAGGTGCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACT	2800					
2810	2820	2830	2840	2850	2860	2870
GCACCTTCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCGGTTCCC	2870					
CGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTGGTCAACCTGCTG	2940					
GTCGCCATGTTTGGCTACACGGTGGGCACCTCCAGGAGAACAATGACCAGGTCTGGAAGTCCAGAGGT	3010					
ACTTCTTGGTGCAGGAGTACGTCAGCGCGCTCAATATCCCCCTTCCCCCTCAGCTCTTGGCTTACTTCTA	3080					
CATGGTGGTGAAGAAGTGGTCAAGTGTGGTGGTGAAGGAGAAACATGGAGTCTTCTGTCTGTCTGTTTC	3150					
3160	3170	3180	3190	3200	3210	3220
AAAAATGAAGACAATGAGACTCTGGCATGGGAGGGTGTGATGAAGGAAAACCTACCTTGTCAGATCAACA	3220					
CAAAAGCCACGACACCTCAGAGGAAATGAGGCACTGGATTTAGACCACTGGATACAAAGCTTAATGATCT	3290					
CAAGGGTCTTCTGAAGAGATTGCTAATAAAATCAAAATAAACTGTATGAACCTCTAATGGAAGAAAATC	3360					
TAATTATAGCAAGATCATATTAAGGAATGCTGATGAACATTTTGGTATCGACTACTAAAAGAGAGATT	3430					
TCAGACCCCTGGGTACATGGTGGATGATTTAAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGT	3500					
3510	3520	3530	3540	3550	3560	3570
GTGATTGGTTCATACCTTGAAGACGGATATAAAGGAAGAAATTTTCTTTATGTGTTCTCCAGAATGGT	3570					
GGCTGTTTCTCTCTGTCTCTCAATGGCTGGGACTGGAGGTGAATGTTTAAAGTGTGTTCTTACCGCCTCC	3640					
TTTTTCTTTAATCTTATTTTGTATGAACACAATATAGGAGAACATCTATCCTATGAATAAGAACCTGG	3710					
TCATGCTTACTCCTGTATTGTATTTGTCTCATTTCCAAATGATTCTCTACTTTTCCCTTTTGTATT	3780					
ATGTGACTAATAGTTGGCATATTGTAAAAAGTCTCTCAAAATAGGCCAGATTCTAAAACATGCTGCAGC	3850					
3860	3870	3880	3890	3900	3910	3920
AAGAGGACCCCGCTCTCTTCAAGGAAAAGTGTTCATTTCTCAGGATGCTTCTTACCTGTGAGAGGAGGT	3920					
GACAAGGCAGTCTCTTGTCTCTTGGACTCACCAGGCTCCTATTGAAGGAACACCCCCATTCTAAATA	3990					
TGTGAAGAGTCCCCAAAATGCAACCTTGAAGGCACTACTGACTTTCTTCTTATTGGATACTCCTCTTA	4060					
TTTATTATTTTCCATTAAAAAATAAGGTGGCTATTAGAAAATTTAGACCATACAGAGATGTAGAAA	4130					
GAACATAAATGTCCCCATTACCTTAAGGTAACTACTGTAACAATTTCTGGATGGTTTTTCAAGCTAT	4200					
4210	4220	4230	4240	4250	4260	4270
TTTTTTCTATGATGTCTCAATCTCTTTCAAAATTTACAGAATGTTATCATACTACATATATACTTT	4270					
TTATGTAAGCTTTTTCACTTAGTATTTTATCAAAATATGTTTTATTATATTATAGCCTTCTTAAACATT	4340					
ATATCAATAATGGCAATAAGGCAACCTCTAGCGATTACATAATTTTGTCTATTGAAGGCTATCTCCAG	4410					
TTGATCATTTGGATGAGCATCTTGTGCAATGAATCCTATTGCTGTATTGGGAAAATTTTCCAGGTTAG	4480					
ATTCCAAATAATATCTATTTATTATTAATAATAAATATCGATTATTTATTAACCAATTTATAAGCT	4550					

Fig. 12A(2)

009060" 6225960

```

      10      20      30      40      50      60      70
+-----+
MRNRRNDTLOSTRILYSSASRSTOLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKCGYAQSQHME 70
GTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYELLTOHWHLKTPNLVISVT 140
GGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLTKYIGEVVRONTISRSSEENIVAIGIAAWGM 210
VSNRDTLIRNCOAEGYFLAQYLMDDFTRDPLY:LONNHTHLLLVDNGCHGHPTVEAKLRNQLKHSERT 280
IQDSNYGGKIPIVCFAOGGGKETLKAINTS[KNK:PCYVVEGSGRIADVIASLVEVEDAPTSSAVKEKLV 350
      360      370      380      390      400      410      420
+-----+
RFLPRTVSRLSEEETESWIKWLKEILECSHLLTV:KMEZAGDEIVSNAISYALYKAFSTSEQQKONWNGQ 420
LKLLLEWNQLDLANDEIFTNDRRWESADLQEVMTALIKDRPKFYRLFLNGLNLRKFLTHDVLTELSN 490
HESTLVYRNLGIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGROEMOIELHGVSPITRHPLQALF[WAI 560
LQNKKELSKVIWECTRGCTLAALGASKLLKTLAKYKNDINAAGESEELANFYETRAVELFTECYSSOEDL 630
AEQLLVYSCEAWGGSNCLELAVEATDQHFTAQPGVONFLSKQWYGEISROTKNWKIILCLFIIPLVGCGF 700
      710      720      730      740      750      760      770
+-----+
VSFRKKPVQKHKKLLWYYVAFFTSPFVYFSWNVVFYIAFLLLFAYVLLMDFHSVPHPPPELVLYSLVFVLF 770
CDEVROWYVNGVNYFTDLWNVMDTLGLEFYIAGIVFRLEHSSNKSSLYSGRV:FCLOY:IFTLRLIHIFTV 840
SRNLGPKIIMLQRMLOVFFFLFLFAYWMVAFGVARGGILRONEQRWRWIFRSVIYEPYLA MFGQVPSDV 910
DGTTYDFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLYCIYMLSTNILLVNLVAMFGYTVGTVDEN 980
NOCVWKFGQRYFLVQEYCSRLNIPFPFIVFAYFMVVKKCFKCCCKEKNMESSVCCFKNEDNETLAWEGYM 1050
      1060      1070      1080      1090      1100      1110      1120
+-----+
KENYLVK:NTKANOTSEEMRHRFRQDQTKLNDLKGLLKE:ANKIK. 1096

```

Fig. 12B

<120> COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

<140> US

<141> 2000-09-06

<160> 877

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 814

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (814)$

<223> n = A, T, C or G

<400> 1

tttttttttt	tttttcacag	tataacagct	ctttattttct	gtgagttcta	ctaggaaatc	60
atcaaatctg	agggttgct	ggaggacttc	aatacacctc	cccccatagt	gaatcagctt	120
ccaggggggc	cagtccctct	ccttacttca	tccccatccc	atgccaaagg	aagaccctcc	180
ctccttggtc	cacagccttc	tctaggtctc	ccagtgcctc	caggacagag	tgggttatgt	240
tttcagctcc	atccttgctg	tgagtgtctg	gtgcgttgty	cctccagctt	ctgctcagtg	300
cttcatggac	agtgtccagc	acatgtcact	ctccactctc	tcagtgtgga	tccactagtt	360
ctagagcggc	cgcacccggg	gtggagctcc	agcttttgtt	ccctttagtg	agggttaatt	420
gcgcgcttgq	cgtaatcatg	gtcataactg	tttctgtgt	gaaattgtta	tccqctcaca	480


```
<220>
<221> misc_feature
<222> (1)...(799)
<223> n = A,T,C or G
```

```
<210> 9
<211> 801
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(801)
<223> n = A,T,C or G
```

```
<210> 10
<211> 789
```

<212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(789)
 <223> n = A,T,C or G

<400> 10
 cagtctatnt ggccagtgtg gcagctttcc ctgtggctgc cgggtgccaca tgccctgtccc 60
 acagtgtggc cgtgggtgaca gcttcagccg ccctcaccgg gttcaccttc tcagccctgc 120
 agatccctgcc ctacacactg gcctccctct accaccggga gaagcagggtg ttcttgccca 180
 aataccgagg ggacactgga ggtgctagca gtgaggacag cctgatgacc agcttcctgc 240
 caggccctaa gcctggagct cccttcctta atggacacgt ggggtgctgga ggcagtggcc 300
 tgctcccacc tccaccgcg ctctgcgggg cctctgcctg tgatgtctcc gtacgtgtgg 360
 tgggtgggtga gccaccgan gccagggtgg ttccgggccc gggcatctgc ctggacctcg 420
 ccatcctgga tagtgcttcc tgctgtccca ngtggcccca tccctgttta tgggctccat 480
 tgtccagctc agccagtctg tcaactgcta tatggtgtct gccgcaggcc tgggtctggt 540
 cccatttact ttgctacaca ggtantattt gacaagaacg anttggccaa atactcagcg 600
 ttaaaaaatt ccagcaacat tgggggtgga aggcctgcct cactgggtcc aactccccgc 660
 tcctgttaac cccatggggc tgccggcttg gccgccaat tctgttgctg ccaaantnat 720
 gtggctctct gctgccacct gttgctggct gaagtgenta cngcncanct nggggggtng 780
 ggngttccc 789

<210> 11
 <211> 772
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(772)
 <223> n = A,T,C or G

<400> 11
 cccaccctac ccaaataatta gacaccaaca cagaaaagct agcaatggat tcccttctac 60
 tttgttaaata aaataagtta aatattttaa tgccctgtgtc tctgtgatgg caacagaagg 120
 accaacaggc cacatcctga taaaaggtaa gaggggggtg gatcagcaaa aagacagtgc 180
 tgtgggctga ggggacctgg ttcttgtgtg ttgcccctca ggactcttcc cctacaaata 240
 actttcatat gttcaaatcc catggaggag tgtttcatcc tagaaactcc catgcaagag 300
 ctacattaaa cgaagctgca ggttaagggg cttanagatg ggaaaccagg tgactgagtt 360
 tattcagctc ccaaaaaccc ttctctaggt gtgtctcaac taggaggcta gctgttaacc 420
 ctgagcctgg gtaatccacc tgcagagtc ccgcattcca gtgcatggaa ccttctggc 480
 ctccctgtat aagtccagac tgaaaccccc ttggaaggnc tccagtcagg cagccctana 540
 aactggggaa aaaagaaaaa gacgccccan ccccagctg tgcanctacg cacctcaaca 600
 gcacaggggtg gcagcaaaaa aaccacttta ctttggcaca aacaaaaact ngggggggca 660
 accccggcac cccnangggg gttaacagga ancngggnaa cntggaacct aattnaggca 720
 ggccnccac cccnaatntt gctgggaaat ttttctccc ctaaattntt tc 772

<210> 12
 <211> 751
 <212> DNA
 <213> Homo sapien

<400> 12

```
<210> 13
<211> 729
<212> DNA
<213> Homo sapien
```

<400> 13

```
<210> 14
<211> 816
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(816)
<223> n = A,T,C or G
```



```

gagcctctgt tagtggagga agattccggg cttcagctaa gtagtcagcg tatgtcccat 240
aagcaaacac tgtgagcagc cggaaggtag aggcaaagtc actctcagcc agctctctaa 300
cattgggcat gtccagcagt tctccaaaca cgtagacacc agnggcctcc agcacctgat 360
ggatgagtggt ggccagcgct gcccccttgg ccgacttggc taggagcaga aattgctcct 420
ggttctgccc tgtcaccttc acttccgcac tcatcactgc actgagtgtg ggggacttgg 480
gctcaggatg tccagagacg tggttccgcc ccctcnctta atgacaccgn ccanncaacc 540
gtcggctccc gccgantgng ttcgtcgtnc ctgggtcagg gtctgctggc cncacttgc 600
aancttcgtc nggcccattg aattcacenc accggaactn gtangatcca ctntttctat 660
aaccggnccg caccgcnntt ggaactccac tcttnttnc tttacttgag ggtaaggtc 720
acccttncg ttaccttggg ccaaaccntn ccntgtgtcg anatngtnaa tcnggncna 780
tnccanccnc atangaagcg ng 802

```

<210> 19

<211> 731

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(731)

<223> n = A,T,C or G

<400> 19

```

cnaagcttcc aggtnacggg ccgcnaancc tgaccnagg tancanaang cagnncggg 60
gagcccaccg tcacngngng gngtctttat nggagggggc ggagccacat cnetggacnt 120
cntgacccca actcccncc nncantgca gtgatgagtg cagaactgaa ggtnacgtgg 180
caggaaccaa gancaaannc tgctccnntc caagtcggcn nagggggcgg ggctggccac 240
gencatecnt cnagtgtctn aaagcccn cctgtctact tgtttggaga acngcnnga 300
catgcccagn gttanataac nggcnagag tnannttgcc tctcccttcc ggctgcgan 360
cngntntgct tagnggacat aacctgacta cttaactgaa ccnngaate tncnccct 420
ccactaaget cagaacaaaa aacttcgaca ccaactcant gtcacctgnc tgcctaagta 480
aagtgtaccc catncccaat gtntgctnga ngctctgncc tgcnttangt tcggctctgg 540
gaagacctat caattnaagc tatgtttctg actgcctctt gctccctgna acaancnacc 600
cnnnntcca agggggggnc ggcccccaat ccccccaacc ntnaattnan tttancccn 660
ccccnggcc cggcctttta cnancntcn nnaacnggna aaaccnnngc tttncccaac 720
nnaatccncc t 731

```

<210> 20

<211> 754

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(754)

<223> n = A,T,C or G

<400> 20

```

tttttttttt tttttttttt taaaaacccc ctccattnaa tgnaaacttc cgaaattgtc 60
caacccctc ntccaaatnn cnttttccgg gnggggggtc caaacccaan ttanntttgg 120
annttaatt aaatnttnnt tggnggnna anccnaatgt nangaaagtt naaccanta 180
tnancttnaa tncctggaaa cngtngntt ccaaaaatnt ttaaccctta antccctccg 240
aatngttna nggaaaaccc aanttctcnt aagggtgtt gaaggntnaa tnaaaanccc 300
nnccaattgt ttttngccac gcctgaatta attggnntcc gntgttttcc nttaaaanaa 360

```



```

nccctcncnc ngncgnannc ctcncncnc gtctcannca ccaccccgcc ccgccaggcc 660
ntcanccacn ggngacnng nagnncnntc gcnccgcgc gcnncnccct cgcncngaa 720
ctnctcngg ccantnncgc tcaanccna cnaaacgcgc ctgcgcggcc cgnagcgncc 780
ncctcncga gtctccccgn ctcccnaccc angnnttcn cgaggacacn nnaccccgcc 840
nncangcgg 849

```

```

<210> 23
<211> 872
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(872)
<223> n = A,T,C or G

```

```

<400> 23
gcgcaaacta tacttcgctc gnactcgtgc gcctcgtcnc tcttttctct ccgaaccatg 60
tctgacnanc ccgattnggc ngatatcnan aagntcganc agtccaaact gantaacaca 120
cacacnncan aganaaatcc nctgccttcc anagtanacn attgaacnng agaaccangc 180
nggcgaatcg taatnaggcg tgcgcgcgca atntgtcncc gtttatntn ccagctcnc 240
ctnccnacc tacntcttcn nagctgtcnn acccctngtn cgnaccccc naggtcggga 300
tcgggtttnn nntgaccgng cnnccctcc cccctccat naeganccnc ccgcaccacc 360
nanngcncgc nccccgncct ctgcgcnc cctgtctntn cccctgtngc ctggcncngn 420
accgcattga cctcgcncn ctncnngaaa ncgnanacgt ccgggttggn annancgctg 480
tgggnnngcg tctgcncgc gtctcttcn ncnncttcca ccatcttont tacngggtct 540
ccnccgctc tcnnncacnc cctgggaagc tntcctntgc ccccttnac tccccctt 600
cgnctgncc cgnccccacc ntcatttnca nacgntcttc acaannncc ggntnncctc 660
cnancngncn gtcancnag ggaagggngg ggnncnntg nttgacgttg ngngangtc 720
cgaanantcc tcncntcan cctaccctc cgggcgncct ctengttnc aacttancaa 780
ntctccccg ngngcncntc tcagcctcnc cnccccnct ctctgcantg tncctctctc 840
tnaccmntac gantnttcn cncctcttt cc 872

```

```

<210> 24
<211> 815
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(815)
<223> n = A,T,C or G

```

```

<400> 24
gcatgcaagc ttgagtattc tatagngtca cctaaatanc ttggcntaat catggtcnta 60
nctgncttcc tgtgtcaaata gtatacnaa tanatatgaa tctnatntga caaganngta 120
tctnccatta gtaacaantg tntgtccat cctgtengan canattccca tnnattncgn 180
cgcattcncn gncantatn taatngggaa ntcnnntnnn ncaccnncat ctatctncc 240
gcnccctgac tggagagat ggatnanttc tntntgacc nacatgttca tcttgattn 300
aanaccccc cgcngnccac cgggtngng cngccnntc ccaagacctc ctgtggaggt 360
aacctgcgtc aganncatca aacntgggaa acccgcncc angtnnaagt ngnnncanan 420
gatcccgctc aggnntnacc atcccttenc agcgcgccct ttngtgcctt anagnnagc 480
gtgtccnanc cctcaacat ganacgcgcc agnccanccg caattnggca caatgtcgn 540
gaacccccca gggggantna tncaaanccc caggattgtc cncncangaa atccncanc 600

```


tccaacencgc gntggccntn cccccccnnn tccttttcccc

820

<210> 27
<211> 818
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(818)
<223> n = A,T,C or G

<400> 27

tctgggtgat	ggcctcttcc	tcctcagga	cctctgactg	ctctgggcca	aagaatctct	60
tgtttcttct	ccgagcccca	ggcagcgggtg	attcagccct	gccaacctg	attctgatga	120
ctgcggtatgc	tgtgacggac	ccaaggggca	aataggggtcc	caggggtccag	ggagggggcgc	180
ctgctgagca	cttccgcccc	tcacctgccc	cagccccctgc	catgagctct	gggctgggtc	240
tccgcctcca	gggttctgct	cttccangca	ngccancaag	tggcgctggg	ccacactggc	300
ttcttctctgc	ccntccctg	gctctganc	tctgtcttcc	tgtcctgtgc	angcnccttg	360
gatctcagtt	tcctcncctc	anngaactct	gtttctgann	tcttcantta	actntgantt	420
tatnaccnan	tggnetgtnc	tgtcnnactt	taatgggccc	gaccggctaa	tccttccctc	480
ntcccttcc	antcnnnna	accngettnc	cntctctcc	ccntancccg	ccngggaanc	540
ctcctttgcc	ctnaccangg	gccnnnaccg	ccctnnctn	ggggggcngg	gtnnctncnc	600
ctgntnnccc	cncctcncnt	tnctctgccc	cnnncnccg	nnccannctc	ncngtccenn	660
tnnctcttcn	ngntcgnaa	ngntcncntn	tnnnnnngnc	ngntnnctnc	tcctctcnc	720
cnnntgnang	tnnttnnnnc	ncngnncccc	nnnnccnnnn	nggnntnnnn	tctncncngc	780
ccnncccc	ngnattaagg	cctccnntct	ccggccnc			818

<210> 28
<211> 731
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

<400> 28

aggaagggcg	gagggatatt	gtangggatt	gagggatagg	agnataangg	gggaggtgtg	60
tccaacacatg	anggtgnngt	tctcttttga	angaggggtg	ngtttttann	ccnggtgggt	120
gattnaaccc	cattgtatgg	agnnaaagg	tttnagggat	ttttcggtc	ttatcagtat	180
ntanattcct	gtnaatcgga	aaatnatntt	tcnnccggaa	aatnttgctc	ccatccgnaa	240
attnctcccc	gtagtgcat	nttngggggg	cngccangtt	tcccaggctg	ctanaatcgt	300
actaaagntt	naagtgggan	tncaaataaa	aacctnncc	agagnatccn	taccgcactg	360
tnnnctnct	tcgcccctng	actctgcngg	agcccaatac	ccnngngnat	gtcncncngn	420
nnngcgncc	tgaaannnn	tcgnggctnn	gancatcang	gggtttcgca	tcaaaagcnn	480
cgtttcncat	naaggcaatt	tngcctcacc	caaccnctng	ccctcnncca	tttngccgtc	540
nggttcncct	acgctnnctg	cncctnnntn	ganattttnc	ccgcctnggg	naancctcct	600
gnaatgggta	gggncttntc	ttttnaccnn	gnggtntact	aatcnnctnc	acgctncttt	660
tctcnacccc	cccccttttt	caatcccanc	ggcnaatggg	gtctccccnn	cgangggggg	720
nnnccccann	c					731

<210> 29

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(799)

<223> n = A,T,C or G

<400> 31

tttttttttt	tttttttggc	gatgctactg	tttaattgca	ggaggtgggg	gtgtgtgtac	60
catgtaccag	ggctattaga	agcaagaagg	aaggagggag	ggcagagcgc	cctgctgagc	120
aacaaaggac	tcctgcagcc	ttctctgtct	gtctcttggc	gcaggcacat	ggggaggcct	180
cccgagggt	gggggccacc	agtccagggt	tgggagcact	acanggggtg	ggagtgggtg	240
gtggctggtg	cnaatggcct	gncacanatc	cctacgattc	ttgacacctg	gatttcacca	300
ggggaccttc	tgttctccca	nggnaacttc	ntnnatctcn	aaagaacaca	actgtttctt	360
cngcanttct	ggctgttcat	ggaaagcaca	ggtgtccnat	ttnggctggg	acttggtaca	420
tatggttccg	gccacctct	ccntcnaan	aagtaattca	ccccccccc	ccntctnttg	480
cctgggccct	taantacca	caccggaact	canttanta	ttcatcttng	gntgggcttg	540
ntnatcnccn	cctgaangcg	ccaagttgaa	aggccacgcc	gtncnccctc	cccatagnan	600
nttttnnct	canctaata	ccccccnggc	aacnatccaa	ttccccccc	tggggggccc	660
agcccanggc	ccccgnetcg	ggnnnccngn	cncgnantcc	ccaggntctc	ccantcngnc	720
ccnnngcncc	ccgcacgcga	gaacanaagg	ntngagccnc	cgcannnnnn	nggtnncnac	780
ctgccccccc	ccnnccngng					799

<210> 32

<211> 789

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(789)

<223> n = A,T,C or G

<400> 32

tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	60
tttttncnag	ggcagggttta	ttgacaacct	cncgggacac	aancaggctg	gggacaggac	120
ggcaacaggc	tcggcgggcg	gcgcgggcgg	ccctacctgc	ggtacccaaat	ntgcagcctc	180
cgtccccgct	tgatnttccct	ctgcagctgc	aggatgccnt	aaaacagggc	ctcggccntn	240
ggtgggcacc	ctgggatttn	aatttccacg	ggcacaatgc	ggtcgcanc	cctcaccacc	300
nattaggaat	agtggtnnta	ccnccnccg	ttggcncact	ccccntggaa	accacttntc	360
gcggtccggg	catctggtct	taaaccttgc	aaacnctggg	gccctctttt	tggttantnt	420
ncnngccaca	atcatnactc	agactggcnc	gggctggccc	caaaaaancn	ccccaaaacc	480
ggnccatgtc	ttnnccgggt	tgctgcnatn	tncatcacct	cccgggcnca	ncaggncaac	540
ccaaaagttc	ttngggcccn	caaaaaanct	ccggggggnc	ccagtttcaa	caaagtcac	600
ccccctggcc	cccaaactct	ccccccgntt	nctgggtttg	ggaacccacg	cctctnnctt	660
tggnnngcaa	gntggntccc	ccttcggggc	ccgggtgggc	ccnnctctaa	ngaaaaacnc	720
ntcctnnnca	ccatcccccc	nngnnacgnc	tancaangna	tccttttttt	tanaaacggg	780
ccccccnccg						789

<210> 33

<211> 793

<212> DNA

<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(793)
 <223> n = A,T,C or G

<400> 33
 gacagaacat gttggatggt ggagcacctt tctatacgac ttacaggaca gcagatgggg 60
 aattcatggc tgttggagca atanaacccc agttctacga gctgctgatc aaaggacttg 120
 gactaaagtc tgatgaactt cccaatcaga tgagcatgga tgattggcca gaaatgaana 180
 agaagtttgc agatgtattt gcaaagaaga cgaaggcaga gtggtgtcaa atctttgacg 240
 gcacagatgc ctgtgtgact cgggttctga cttttgagga ggttgttcat catgatcaca 300
 acaangaacg gggctcgttt atcaccantg aggagcagga cgtgagcccc cgcctgcac 360
 ctctgctgtt aaacacccca gccatccctt ctttcaaaag ggatccacta cttctagagc 420
 ggncgccacc gcggtggagc tccagctttt gtcccttcta gtgagggtta attgcgcgct 480
 tggcgtaatc atgggtcatan ctgtttcctg tgtgaaattg ttatccgctc acaattccac 540
 acaacatacg anccggaagc atnaaatttt aaagcctggn ggtngcctaa tgantgaact 600
 nactcacatt aattggcttt gcgctcactg cccgctttcc agtccggaaa acctgtcctt 660
 gccagctgcc nttaatgaat cnggccacc cccggggaaa aggcngtttg cttnttgggg 720
 cgcncctccc gctttctcgc ttcctgaant ccttcccccc ggtctttcgg cttgcggcna 780
 acggtatcna cct 793

<210> 34
 <211> 756
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(756)
 <223> n = A,T,C or G

<400> 34
 gccgcgaccg gcatgtacga gcaactcaag ggcgagtgga accgtaaaag ccccaatctt 60
 ancaagtgcg gggaanagct gggtcgactc aagctagtgc tcttgagact caacttcttg 120
 ccaaccacag ggaccaagct gaccaaacag cagctaattc tggcccgtga catactggag 180
 atcggggccc aatggagcat cctacgcaan gacatccctt ctttcgagcg ctacatggcc 240
 cagctcaaat gctactactt tgattacaan gagcagctcc ccgagtcagc ctatatgcac 300
 cagctcttgg gcctcaacct cctcttctctg ctgtcccaga accgggtggc tgantnccac 360
 acgganttgg ancggtgcgc tgcccaanga catacanacc aatgtctaca tcnaccacca 420
 gtgtcctgga gcaatactga tgganggcag ctaccncaa gtnttccttg ccnagggtaa 480
 catccccgcg cgagagctac accttcttca ttgacatcct gctcgacact atcagggatg 540
 aaaatcgng ggttgctcca gaaaggctnc aanaanatcc ttttcnctga aggcccccg 600
 atncnctagt nctagaatcg gcccgccatc gcggtgganc ctccaacctt tcgttnccct 660
 ttactgaggg ttnattgccg cccttggcgt tatcatggct acnccngttn cctgtgttga 720
 aattnttaac ccccacaat tccacgccna cattng 756

<210> 35
 <211> 834
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(834)

<223> n = A,T,C or G

<400> 35

```

ggggatctct anactnacct gnatgcatgg ttgtcgggtg ggctcgtgtc gatgaanatg      60
aacaggatct tgcccttgaa gctctcggct gctgtnttta agttgctcag tctgccgtca      120
tagtcagaca cncctctggg caaaaaacan caggatntga gtcttgattt cacctccaat      180
aatcttcngg gctgtctgct cgggtgaactc gatgacnang ggcagctggg tgtgtntgat      240
aaantccanc angttctcct tggtgacctc cccttcaaag ttgttcgggc cttcatcaaa      300
cttctnnaan angannancc canctttgtc gagctggnat ttgganaaca cgtcactggt      360
ggaaactgat cccaaatggg atgtcatcca tcgcctctgc tgcctgcaaa aaacttgctt      420
ggcncaaate cgactcccn tccttgaaag aagccnatca cccccccctc cctggactcc      480
nncaangact ctncgcctnc ccntccnng cagggttggg ggcannccgg gcccntgcgc      540
ttcttcagcc agttcacnat nttcatcagc ccctctgcca gctgtnttat tccttggggg      600
ggaanccgtc tctcccttcc tgaannaact ttgaccgtng gaatagccgc gentcnent      660
acntnctggg cggggttcaa antccctecn ttgcnntcn cctcgggcca ttctggattt      720
nccnaacttt ttcttcccc cncctcncgg ngtttggntt tttcatnggg ccccaactct      780
gctnttggcc antccctg gggcntntan cncctcctnt ggtcccntng ggcc      834

```

<210> 36

<211> 814

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (814)

<223> n = A,T,C or G

<400> 36

```

cggncgcttt ccngccgcgc ccggtttcca tgacnaaggc tcccttcang ttaaatacnn      60
cctagnaaac attaatgggt tgctctacta atacatcata cnaaccagta agcctgcca      120
naacgccaac tcaggccatt cctaccaaag gaagaaaggc tggctctctc acccctgta      180
ggaaaggcct gccttgtaag acaccacaat ncggctgaat ctnaagtctt gtgttttact      240
aatggaaaaa aaaaataaac aanaggtttt gttctcatgg ctgcccaccg cagcctggca      300
ctaaaacanc ccagcgctca cttctgcttg ganaaatatt ctttgcctt ttggacatca      360
ggcttgatgg tatcactgcc acntttccac ccagctgggc ncccttcccc catntttgtc      420
antganctgg aaggcctgaa ncttagtctc caaaagtctc ngcccacaag accggccacc      480
aggggagngt ntttncagtg gatctgcca anantaccn tatcatcnnt gaataaaaag      540
gcccctgaac ganatgcttc cancanctt taagacccat aatcctngaa ccatggtgcc      600
cttcgggtct gatccnaag gaatgttctt gggctccant ccctcctttg ttncttacgt      660
tgtnttggac cntgtctngn atnaccaan tganatcccc ngaagcacc tncctctggc      720
atgtganttt cntaaattct ctgccctacn nctgaaagca cnattccctn ggcnccnaan      780
ggngaactca agaaggctcn ngaaaaacca cncn      814

```

<210> 37

<211> 760

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (760)

<223> n = A,T,C or G

```

<400> 37
gcagtgctgt cttcctcaaa gttgttcttg ttgccataac aaccaccata ggtaaagcgg      60
gcgcagtggt cgctgaaggg gttgtagtac cagcgcggga tgctctcctt gcagagtcct      120
gtgtctggca ggtccacgca atgcccttg tcaactggga aatggatgcg ctggagctcg      180
tcnaanccac tcgtgtatth ttacangca gcctcctccg aagcntccgg gcagttgggg      240
gtgtcgtcac actccactaa actgtcgatn cancagccca ttgctgcagc ggaactgggt      300
gggctgacag gtgccagaac aactgggatn ggcctttcca tgggaaggcc tgggggaaat      360
cncctnancc caaactgcct ctcaaaggcc accttgaca ccccgacagg ctagaaatgc      420
actcttcttc ccaaaggtag ttgttcttgt tgcccaagca ncctccanca aaccaaanc      480
ttgcaaaatc tgctccgtgg gggcatnnn taccanggtt ggggaaanaa acccgcgngn      540
ganccnctt gtttgaatgc naaggnaata atcctcctgt cttgcttggg tgggaanagca      600
caattgaact gttaacnttg ggccgngttc cctnnggtg gtctgaaact aatcacgcgc      660
actggaaaaa ggtangtgcc ttcttgaat tcccaaantt cccctngntt tgggtntttt      720
ctcctctncc ctaaaaatcg tnttcccccc cntangggcg      760

```

<210> 38

<211> 724

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (724)

<223> n = A,T,C or G

```

<400> 38
tttttttttt tttttttttt tttttttttt tttttaaaaa cccctcccat tgaatgaaaa      60
cttcnnaaat tgtccaaccc cctcnnccaa atnnccattt ccgggggggg gttccaaacc      120
caaattaatt ttgganttta aattaaatnt tnattngggg aanaanccaa atgtnaagaa      180
aatttaaccc attatnaact taaatnctn gaaaccntg gnttccaaaa atttttaacc      240
cttaaattccc tccgaaattg ntaanggaaa accaaattcn cctaaggctn tttgaagggt      300
ngatttaaac ccccttnant tnttttnacc cmngnctnaa ntatttngnt tccggtgttt      360
tcctnttaan cntnggtaac tcccngtaat gaannncctt aanccaatta aaccgaattt      420
tttttgaatt ggaaattccn ngggaattna ccgggggttt tccnttttg gggccatncc      480
ccncttttcg gggtttgggn ntagggtgaa tttttnnang ncccaaaaaa ncccccaana      540
aaaaaactcc caagnnttaa ttngaattnc ccccttccca ggccttttg gaaaggnggg      600
ttnttggggg ccngggantt cnttcccccn ttncncccc ccccccnggt aaanggttat      660
ngnntttggt ttttgggccc cttnanggac ctccggatn gaaattaaat ccccggnccg      720
gccg      724

```

<210> 39

<211> 751

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (751)

<223> n = A,T,C or G

```

<400> 39
tttttttttt tttttctttg ctcacattta atttttatth tgattttttt taatgctgca      60
caacacaata tttatttcat ttgtttcttt tatttcattt tatttgtttg ctgctgctgt      120
tttattttat tttactgaaa gtgagaggga acttttgtgg ctttttttcc tttttctgta      180

```


<400> 42
 acttactgaa tttagttctg tgctcttcct tatttagtgt tgtatcataa atactttgat 60
 gtttcaaaca ttctaaataa ataattttca gtggcttcat a 101

<210> 43
 <211> 305
 <212> DNA
 <213> Homo sapien

<400> 43
 acatctttgt tacagtctaa gatgtgttct taaatcacca ttccttcctg gtcctcaccc 60
 tccagggtgg tctcacactg taattagagc tattgaggag tctttacagc aaattaagat 120
 tcagatgcct tgctaagtct agagttctag agttatgttt cagaaaagtct aagaaaccca 180
 cctcttgaga ggtcagtaaa gaggacttaa tatttcatat ctacaaaatg accacaggat 240
 tggatacaga acgagagtta tcctggataa ctcagagctg agtacctgcc cgggggcccgc 300
 tcgaa 305

<210> 44
 <211> 852
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1) ... (852)
 <223> n = A,T,C or G

<400> 44
 acataaatat cagagaaaag tagtctttga aatatttacg tccaggagtt ctttgtttct 60
 gattatttgg tgtgtgtttt ggtttgtgtc caaagtattg gcagcttcag ttttcatttt 120
 ctctccatcc tcgggcattc ttcccaaatt tatataccag tcttcgtcca tccacacgct 180
 ccagaatttc tctttttag tagtatctca tagctcggt gagcttttca taggtcatgc 240
 tgctgttggt cttcttttta ccccatagct gagccactgc ctctgatttc aagaacctga 300
 agacgcctc agatcgggtc tcccatttta ttaatcctgg gttcttgtct gggttcaaga 360
 ggatgtcgcg gatgaattcc cataagttag tccctctcgg gttgtgcttt ttggtgtggc 420
 acttggcagg ggggtcttgc tcctttttca tatcagggtga ctctgcaaca ggaagggtgac 480
 tgggtggtgt catggagatc tgagcccggc agaaagtatt gctgtccaac aaatctactg 540
 tgctaccata gttggtgtca tataaatagt tctngtcttt ccagggtgtc atgatggaag 600
 gctcagtttg ttcagtcttg acaatgacat tgtgtgtgga ctggaacagg tccactactgc 660
 actggccggt ccacttcaga tgctgcaagt tgctgtagag gagntgcccc gccgtccctg 720
 ccgcccgggt gaactcctgc aaactcatgc tgcaaagggt ctcgccgttg atgtcgaact 780
 cntggaaagg gatacaattg gcatccagct ggttgggtgc caggaggtga tggagccact 840
 cccacacctg gt 852

<210> 45
 <211> 234
 <212> DNA
 <213> Homo sapien

<400> 45
 acaacagacc cttgctcgct aacgacctca tgctcatcaa gttggacgaa tccgtgtccg 60
 agtctgacac catcgggagc atcagcattg cttcgcagtg ccctaccgcg gggaactctt 120
 gcctcgtttc tggctggggc ctgctggcga acggcagaat gcctaccgtg ctgcagtgcg 180
 tgaacgtgtc ggtggtgtct gaggaggtct gcagtaagct ctatgacctg ctgt 234

<210> 46
 <211> 590
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(590)
 <223> n = A,T,C or G

<400> 46
 acttttttatt taaatgttta taaggcagat ctatgagaat gatagaaaac atgggtgtgta 60
 atttgatagc aatatttttg agattacaga gtttttagtaa ttaccaatta cacagttaaa 120
 aagaagataa tatattccaa gcanatacaa aatatctaata gaaagatcaa ggcaggaaaa 180
 tgantataac taattgacaa tggaaaatca attttaatgt gaattgcaca ttatccttta 240
 aaagctttca aaanaaanaa ttattgcagt ctanttaatt caaacagtggt taaatgggtat 300
 caggataaan aactgaaggc canaaagaat taattttcac ttcattgtaac ncacccanac 360
 ttacaatggc ttaaatgcan ggaaaaagca gtgggaagtag ggaagtantc aagggtctttc 420
 tgggtctctaa tctgccttac tctttgggtg tggctttgat cctctggaga cagctgocag 480
 ggctcctgtt atatccacaa tcccagcagc aagatgaagg gatgaaaaag gacacatgct 540
 gccttccttt gaggagactt catctcactg gccaacactc agtcacatgt 590

<210> 47
 <211> 774
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(774)
 <223> n = A,T,C or G

<400> 47
 acaagggggc ataatgaagg agtggggana gatttttaaag aaggaaaaaa aacgaggccc 60
 tgaacagaat ttctcctgnac aacggggcctt caaaataatt ttcttgggga gggttcaagac 120
 gcttctactgc ttgaaactta aatgggatgtg ggacanaatt ttctgtaatg accctgaggg 180
 cattacagac gggactctgg gaggaaggat aaacagaaaag gggacaaaagg ctaatcccaa 240
 aacatcaaag aaaggaaggc ggcgtcatat ctcccagcct acacagttct ccagggtctct 300
 cctcatccct ggaggacgac agtggaggaa caactgacca tgtccccagg ctctgtgtgtg 360
 ctggctcctg gtcttcagcc cccagctctg gaagcccacc ctctgtgtgtg cctgctgtggc 420
 ccacactcct tgaacacaca tcccaggtt atattcctgg acatggctga acctcctatt 480
 cctacttccg agatgccttg ctccctgcag cctgtcaaaa tccactcac cctccaaacc 540
 acggcatggg aagcctttct gacttgcttg attactccag catcttggaa caatccctga 600
 ttcccactc cttagaggca agatagggtg gttaagagta gggctggacc acttgagacc 660
 aggctgctgg cttcaaattt tggctcattt acgagctatg ggaccttggg caagtnatct 720
 tcacttctat gggcctcatt ttgttctacc tgcaaaatgg gggataataa tagt 774

<210> 48
 <211> 124
 <212> DNA
 <213> Homo sapien

<220>

```

<221> misc_feature
<222> (1)...(124)
<223> n = A,T,C or G

<400> 48
canaaattga aattttataa aaaggcattt ttctcttata tccataaaat gatataattt      60
ttgcaantat anaaatgtgt cataaattat aatgttcctt aattacagct caacgcaact      120
tggt                                             124

<210> 49
<211> 147
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(147)
<223> n = A,T,C or G

<400> 49
gccgatgcta ctattttatt gcaggagggtg ggggtgtttt tattattctc tcaacagctt      60
tgtggctaca ggtggtgtct gactgcatna aaaanttttt tacgggtgat tgcaaaaatt      120
ttagggcacc catatcccaa gcantgt                                             147

<210> 50
<211> 107
<212> DNA
<213> Homo sapien

<400> 50
acattaaatt aataaaagga ctgttggggt tctgctaaaa cacatggctt gatatatattgc      60
atggtttgag gttaggagga gttaggcata tgttttggga gaggggt                                             107

<210> 51
<211> 204
<212> DNA
<213> Homo sapien

<400> 51
gtcctaggaa gtctagggga cacacgactc tggggtcacg gggccgacac acttgcacgg      60
cggaaggaa aggcagagaa gtgacaccgt cagggggaaa tgacagaaag gaaaatcaag      120
gccttgcaag gtcagaaagg ggactcaggg cttccaccac agccctgccc cacttggcca      180
cctccctttt gggaccagca atgt                                             204

<210> 52
<211> 491
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G

```


<400> 52
 acaaagataa catttatctt ataacaaaaa tttgatagtt ttaaagggtta gtattgtgta 60
 gggatattttc caaaagacta aagagataac tcaggtaaaaa agttagaaat gtataaaaaca 120
 ccatcagaca gggtttttaa aaacaacata ttacaaaatt agacaatcat ccttaaaaaa 180
 aaaacttctt gtatcaattt cttttgttca aaatgactga cttaantatt tttaaatatt 240
 tcanaaacac ttcctcaaaa attttcaana tggtagcttt canatgtncc ctcagtccca 300
 atgttgctca gataaataaa tctcgtgaga acttaccacc caccacaagc tttctggggc 360
 atgcaacagt gtcttttctt tnttttttct tttttttttt ttacaggcac agaaactcat 420
 caattttatt tggataacaa agggctctcca aattatattg aaaaataaat ccaagttaat 480
 atcactcttg t 491

<210> 53
 <211> 484
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(484)
 <223> n = A,T,C or G

<400> 53
 acataattta gcagggctaa ttaccataag atgctattta ttaanaggtn tatgatctga 60
 gtattaacag ttgctgaagt ttggatattt tatgcagcat tttctttttg ctttgataac 120
 actacagaac ccttaaggac actgaaaatt agtaagtaaa gttcagaaac attagctgct 180
 caatcaaatt cttacataac actatagtaa ttaaaacgtt aaaaaaaagt gttgaaatct 240
 gcactagtat anaccgctcc tgtcaggata anactgcttt ggaacagaaa gggaaaaanc 300
 agctttgant ttctttgtgc tgatangagg aaaggctgaa ttaccttggt gcctctccct 360
 aatgattggc aggtcnggta aatnccaaaa catattccaa ctcaacactt cttttccncc 420
 tancttgant ctgtgtattc caggancagg cggatggaat gggccagccc ncggatgttc 480
 cant 484

<210> 54
 <211> 151
 <212> DNA
 <213> Homo sapien

<400> 54
 actaaacctc gtgcttgtga actccataca gaaaacgggtg ccatccctga acacggctgg 60
 ccactgggta tactgctgac aaccgcaaca acaaaaacac aaatccttgg cactggctag 120
 tctatgtcct ctcaagtgcc tttttgtttg t 151

<210> 55
 <211> 91
 <212> DNA
 <213> Homo sapien

<400> 55
 acctggcttg tctccgggtg gttcccggcg cccccacgg tccccagAAC ggacactttc 60
 gccctccagt ggatactcga gccaaagtgg t 91

<210> 56
 <211> 133
 <212> DNA

<400> 56

<210> 57

<212> DNA

 $\langle 220 \rangle$ $\langle 222 \rangle \quad (1) \dots (147)$

<223> n = A, T, C or G

<400> 57

<210> 58

<211> 198

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (198)$

<223> n = A, T, C or G

<400> 58

<210> 59

<211> 330

<212> DNA

<213> Homo sapien

<400> 59

<210> 60

<211> 175

<212> DNA

<400> 65

```
<210> 66
<211> 305
<212> DNA
<213> Homo sapien
```

```
<210> 67
<211> 385
<212> DNA
<213> Homo sapien
```

```
<210> 68
<211> 73
<212> DNA
<213> Homo sapien
```

```
<210> 69
<211> 536
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (536)
<223> n = A, T, C or G
```

```
<210> 70
<211> 477
<212> DNA
<213> Homo sapien
```

```
<210> 71
<211> 533
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(533)
<223> n = A,T,C or G
```

```
<210> 72
<211> 511
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc feature
```

<222> (1)...(511)

<223> n = A,T,C or G

<400> 72

tattacggaa	aaacacacca	cataattcaa	ctancaaaga	anactgcttc	agggcgtgta	60
aaatgaaagg	cttccaggca	gttatctgat	taaagaacac	taaaagaggg	acaaggctaa	120
aagccgcagg	atgtctacac	tatancaggc	gctatttggg	ttggctggag	gagctgtgga	180
aaacatggan	agattgggtg	tgganacgc	cgtggctatt	cctcattgtt	attacanagt	240
gaggttctct	gtgtgcccac	tggtttgaaa	accgttctnc	aataatgata	gaatagtaca	300
cacatgagaa	ctgaaatggc	ccaaacccag	aaagaaaagg	caactagatc	ctcagaanac	360
gcttctaggg	acaataaccg	atgaagaaaa	gatggcctcc	ttgtgcccc	gtctgttatg	420
atttctctcc	attgcagcna	naaacccgtt	cttctaagca	aacncagggtg	atgatggcna	480
aaatacaccc	cctcttgaag	naccnggagg	a			511

<210> 73

<211> 499

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(499)

<223> n = A,T,C or G

<400> 73

cagtgccagc	actggtgcca	gtaccagtac	caataacagt	gccagtgcc	gtgccagcac	60
cagtgggtgg	ttcagtgctg	gtgccagcct	gaccgccact	ctcacatttg	ggctcttcgc	120
tggccttggg	ggagctgggt	ccagcaccag	tggcagctct	ggtgcctgtg	gtttctccta	180
caagtgagat	tttagatatt	gttaatcctg	ccagtctttc	tcttcaagcc	aggggtgcac	240
ctcagaaacc	tactcaacac	agcactctag	gcagccacta	tcaatcaatt	gaagttgaca	300
ctctgcatta	aatctatttg	ccatttctga	aaaaaaaaaa	aaaaaaaaagg	cggccgctcg	360
antctagagg	gcccgtttta	acccgctgat	cagcctcgac	tgtgccttct	anttgcagc	420
catctgttgt	ttgcccctcc	cccgntgcct	tccttgaccc	tggaaagtgc	cactcccaact	480
gtccttttct	aantaaat					499

<210> 74

<211> 537

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(537)

<223> n = A,T,C or G

<400> 74

tttcatagga	gaacacactg	aggagatact	tgaagaattt	ggattcagcc	gcgaagagat	60
ttatcagctt	aactcagata	aatcattga	aagtaataag	gtaaaagcta	gtctctaact	120
tccaggccca	cggctcaagt	gaatttgaat	actgcattta	cagtgtagag	taacacataa	180
cattgtatgc	atggaaacat	ggaggaacag	tattacagtg	tcctaccact	ctaatacaaga	240
aaagaattac	agactctgat	tctacagtga	tgattgaatt	ctaaaaatgg	taatcattag	300
ggcttttgat	ttataanact	ttgggtactt	atactaaatt	atggtagtta	tactgccttc	360
cagtttgctt	gatataattg	ttgatattaa	gattccttgac	ttatattttg	aatgggttct	420
actgaaaaan	gaatgatata	ttcttgaaga	catcgatata	cattttattta	cactcttgat	480

tctacaatgt agaaaatgaa ggaaatgccc caaattgtat ggtgataaaa gtcccgt 537

<210> 75
<211> 467
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(467)
<223> n = A,T,C or G

<400> 75
caaanacaat tgttcaaaag atgcaaatac tacactactg ctgcagctca caaacacctc 60
tgcataattac acgtacctcc tctgtctcct caagtagtgt ggtctatatt gccatcatca 120
cctgctgtct gcttagaaga acggctttct gctgcaangg agagaaatca taacagacgg 180
tggcacaagg aggccatctt ttctcatcgc gttattgtcc ctagaagcgt cttctgagga 240
tctagtggg ctttctttct gggtttgggc catttcantt ctcatgtgtg tactattcta 300
tcattattgt ataacggttt tcaaaccngt gggcacncag agaacctcac tctgtaataa 360
caatgaggaa tagccacggt gatctccagc accaaatctc tccatgttnt tccagagctc 420
ctccagccaa cccaaatagc cgctgctatn gtgtagaaca tccctgn 467

<210> 76
<211> 400
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

<400> 76
aagctgacag cattcgggcc gagatgtctc gctccgtggc cttagctgtg ctgcgcctac 60
tctctctttc tggcctggag gctatccagc gtactccaaa gattcaggtt tactcacgtc 120
atccagcaga gaatggaaag tcaaatttcc tgaattgcta tgtgtctggg ttcatccat 180
ccgacattga agttgactta ctgaagaatg gagagagaat tgaaaaagtg gagcattcag 240
acttgtcttt cagcaaggac tggcttttct atctcttgta ctacactgaa ttcccccca 300
ctgaaaaaga tgagtatgcc tgccgtgtga accatgtgac tttgtcacag cccaagatng 360
ttnagtggga tcganacatg taagcagcan catgggaggt 400

<210> 77
<211> 248
<212> DNA
<213> Homo sapien

<400> 77
ctggagtgcc ttggtgtttc aagccccctgc aggaagcaga atgcaccttc tgaggcacct 60
ccagctgccc cggcggggga tgcgaggctc ggagcaccct tgccgggtg tgattgtctc 120
caggcactgt tcatctcagc ttttctgtcc ctttgtcccc ggcaagcgt tctgtgaaa 180
gttcatactt ggagcctgat gtcttaacga ataaaggtcc catgctccac ccgaaaaaaa 240
aaaaaaaaa 248

<210> 78

<400> 78

<400> 79

<400> 80

```
<210> 81
<211> 232
```


<210> 84
 <211> 380
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(380)
 <223> n = A,T,C or G

<400> 84
 gctggtagcc tatggcgtgg ccacggagg gtcctgagg cacgggacag tgacttccca 60
 agtatacctgc gccgcgtctt ctaccgtccc tacctgcaga tcttcgggca gattccccag 120
 gaggacatgg acgtggccct catggagcac agcaactgct cgtcggagcc cggcttcttg 180
 gcacaccctc ctggggccca ggccggcacc tgcgtctccc agtatgccaa ctggctggtg 240
 gtgctgctcc tcgtcatctt cctgctcgtg gccaacatcc tgctggtcac ttgctcattg 300
 ccatgttcag ttacacattc ggcaaagtac agggcaacag cnatctctac tgggaaggcc 360
 agcgttnccg cctcatccgg 380

<210> 85
 <211> 481
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(481)
 <223> n = A,T,C or G

<400> 85
 gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggctctctgc ttcataccgc 60
 tnccatcgtc atactgtagg tttgccacca cctcctgcat cttggggcgg ctaatatcca 120
 ggaaactctc aatcaagtca ccgtcnatna aacctgtggc tggttctgtc ttcgcctcgg 180
 tgtgaaagga tctccagaag gagtgcctga tcttccccac acttttgatg actttattga 240
 gtogattctg catgtccagc aggaggttgt accagctctc tgacagttag gtcaccagcc 300
 ctatcatgcc nttgaacgtg ccgaagaaca ccgagccttg tgtggggggg gnagtctcac 360
 ccagattctg cattaccaga nagccgtggc aaaaganatt gacaactcgc ccaggngaa 420
 aaagaacacc tcttggaagt gctngccgct cctcgtccnt tgggtggnngc gentnecctt 480
 t 481

<210> 86
 <211> 472
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(472)
 <223> n = A,T,C or G

<400> 86
 aacatcttcc tgtataatgc tgtgtaatat cgatccgatn ttgtctgctg agaattcatt 60
 acttggaana gcaacttnaa gcctggacac tggattataa attcacaata tgcaaacatt 120
 taaacagtgt gtcaatctgc tcccttactt tgtcatcacc agtctgggaa taagggtatg 180

```
<210> 87
<211> 413
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(413)
<223> n = A,T,C or G
```

```
<210> 88
<211> 448
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(448)
<223> n = A,T,C or G
```

```
<210> 89
<211> 463
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(463)
<223> n = A,T,C or G
```

<400> 89
gaattttgtg cactggccac tgtgatggaa ccattggggcc aggatgcttt gagtttatca 60
gtagtgattc tgccaaagtt ggtgttgtaa catgagtatg taaaatgtca aaaaattagc 120
agaggtctag gtctgcatat cagcagacag tttgtccgtg tattttgtag cettgaagtt 180
ctcagtgaca agttnnttct gatgcgaagt tctnattcca gtgttttagt cetttgcac 240
tttnatgttn agacttgccct ctntnaaatt gcttttgtnt tctgcaggta ctatctgtgg 300
tttaacaaaa tagaannact tctctgcttn gaanatttga atatcttaca tctnaaaatn 360
aattctctcc ccatannaaa acccangccc ttggganaat ttgaaaaang gntccttcnn 420
aattcnana anttcagntn tcatacaaca naacngganc ccc 463

<210> 90
<211> 400
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

<400> 90
agggattgaa ggtctntnt actgtcggac tgttcancca ccaactctac aagttgctgt 60
cttccactca ctgtctgtaa gcntnttaac ccagactgta tcttcataaa tagaacaat 120
tcttcaccag tcacatcttc taggaccttt ttggattcag ttagtataag ctcttccact 180
tcctttgtta agacttcac tggtaaagtc ttaagttttg tagaaaggaa ttttaattgct 240
cgttctctaa caatgtctc tccttgaagt atttggctga acaacccacc tnaagtcct 300
ttgtgcatcc attttaaata tacttaatag ggcattggtg cactagggtta aattctgcaa 360
gagtcactctg tctgcaaaag ttgcgttagt atatctgcc 400

<210> 91
<211> 480
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(480)
<223> n = A,T,C or G

<400> 91
gagctcggat ccaataatct ttgtctgagg gcagcacaca tatncagtgc catggnaact 60
ggtctacccc acatgggagc agcatgccgt agntatataa ggtcattccc tgagtcagac 120
atgectcttt gactaccgtg tgccagtgtt ggtgattctc acacacctcc nnccgctctt 180
tgtggaaaaa ctggcacttg nctggaacta gcaagacatc acttacaat tcacccacga 240
gacacttgaa aggtgtaaca aagcgactct tgcattgctt tttgtccctc cggcaccagt 300
tgtcaatact aaccgcgtgg tttgcctcca tcacatttgt gatctgtagc tctggataga 360
tctcctgaca gtactgaaga acttcttctt ttgtttcaaa agcaactctt ggtgcctgtt 420
ngatcagggtt cccatttccc agtccgaatg ttcacatggc atatnttact tcccacaaaa 480

<210> 92
<211> 477
<212> DNA
<213> Homo sapien

<400> 92

<210> 93

<212> DNA

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (377)$

<223> n = A, T, C or G

<400> 93

<210> 94

<211> 495

<212> DNA

<213> Homo sapien

 $\langle 220 \rangle$

<221> misc feature

<222> (1) ... (495)

<223> n = A, T, C or G

<400> 94

cctttgagg	ggttagggtc	cagttcccag	tggaagaaac	aggccaggag	aantgcgtgc	60
cgagctgang	cagatttccc	acagtgaccc	cagagccctg	ggctatagtc	tctgaccct	120
ccaaggaaag	accaccttct	ggggacatgg	gctggagggc	aggacctaga	ggcaccaagg	180
gaaggcccca	ttccggggct	gttccccgag	gaggaaggga	aggggctctg	tgtgcccccc	240
acgaggaana	ggccctgant	cctgggatca	nacaccctt	cacgtgtatc	cccacacaaa	300
tgcaagctca	ccaagggtccc	ctctcagtc	cttccctaca	ccctgaacgg	ncactggccc	360
acaccacccc	agancancca	ccgcctatgg	ggaatgtnct	caaggaatcg	cngggcaacg	420
tggactctng	tcccnnaagg	gggcagaatc	tccaatagan	gganngaacc	cttgctnana	480

aaaaaaaaana aaaaaa

495

<210> 95
 <211> 472
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(472)
 <223> n = A,T,C or G

<400> 95
 gggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc 60
 cctctggaag ccttgccgag agcggacttt gtaattgttg gagaataact gctgaatttt 120
 tagctgtttt gagttgattc gcaccactgc accacaactc aatatgaaaa ctatttnact 180
 tattttattat cttgtgaaaa gtatacaatg aaaattttgt tcatactgta tttatcaagt 240
 atgatgaaaa gcaatagata tatattcttt tattatgttn aattatgatt gccattatta 300
 atcggcaaaa tgtggagtgat atgttctttt cacagtaata tatgcctttt gtaacttcac 360
 ttgggtattt tattgtaaat gaattacaaa attcttaatt taagaaaatg gtangttata 420
 tttanttcan taatttcttt ccttgtttac gttaattttg aaaagaatgc at 472

<210> 96
 <211> 476
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(476)
 <223> n = A,T,C or G

<400> 96
 ctgaagcatt tcttcaaact tntctacttt tgtcattgat acctgtagta agttgacaat 60
 gtggtgaaat ttcaaaaatta tatgtaactt ctactagttt tactttctcc cccaagtctt 120
 ttttaactca tgattttttac acacacaatc cagaacttat tatatagcct ctaagtcttt 180
 attcttcaca gtagatgatg aaagagtccct ccagtgtctt gngcanaatg ttctagntat 240
 agctggatac atacngtggg agttctataa actcatacct cagtgggact naacccaaat 300
 tgtgttagtc tcaattccta ccacactgag ggagcctccc aaatcactat attcttatct 360
 gcaggtactc ctccagaaaa acngacaggg caggcttgca tgaaaaagtn acatctgcgt 420
 taaaaagtct atcttctctca nangtctgtn aaggaacaat ttaatcttct agcttt 476

<210> 97
 <211> 479
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(479)
 <223> n = A,T,C or G

<400> 97
 actcttttcta atgctgatat gatcttgagt ataagaatgc atatgtcact agaatggata 60

aaataatgct	gcaaaacttaa	tgttcttatg	caaaatggaa	cgctaataa	acacagctta	120
caatcgcaaa	tcaaaaactca	caagtgtctca	tctgttgtag	atttagtgta	ataagactta	180
gattgtgctc	cttcggatat	gattgtttct	canatcttgg	gcaatnttcc	ttagtcaa	240
caggctacta	gaattctgtt	attggatatn	tgagagcatg	aaatTTTTaa	naatacactt	300
gtgattatna	aattaatcac	aaattttact	tatacctgct	atcagcagct	agaaaaacat	360
ntnnttttta	natcaaagta	ttttgtgttt	ggaantgtnn	aaatgaaatc	tgaatgtggg	420
ttcnatctta	ttttttcccn	gacnactant	tnctttttta	gggnctattc	tgancctatc	479

<210> 98

<211> 461

<212> DNA

<213> Homo sapien

<400> 98

agtgacttgt	cctccaacaa	aacccttga	tcaagtttgt	ggcactgaca	atcagaccta	60
tgctagttcc	tgtcatctat	tcgctactaa	atgcagactg	gaggggacca	aaaaggggca	120
tcaactccag	ctggattatt	ttggagcctg	caaactctatt	cctacttgta	cggactttga	180
agtgattcag	tttctctac	ggatgagaga	ctggctcaag	aatatcctca	tcagacttta	240
tgaagccact	ctgaacacgc	tgttatcta	gatgagaaca	gagaaataaa	gtcagaaaat	300
ttacctggag	aaaagaggct	ttggctgggg	accatcccat	tgaaccttct	cttaaggact	360
ttaagaaaaa	ctaccacatg	ttgtgtatcc	tgggtgccggc	cgtttatgaa	ctgaccaccc	420
tttgaataaa	tcttgacgct	cctgaacttg	ctcctctgcg	a		461

<210> 99

<211> 171

<212> DNA

<213> Homo sapien

<400> 99

gtggcgcgcg	gcaggtgttt	cctcgtaccg	cagggccccc	tcccttcccc	aggcgctccct	60
cggcgccctct	gcgggcccga	ggaggagcgg	ctggcggtgtg	gggggagtgt	gacccaccct	120
cggtgagaaa	agccttctct	agcgatctga	gaggcgtgcc	ttgggggtac	c	171

<210> 100

<211> 269

<212> DNA

<213> Homo sapien

<400> 100

cggcgcgaag	tgcaactcca	gctggggccg	tgcgacgaa	gattctgcca	gcagttggtc	60
cgactgcgac	gacggcggcg	gcgacagtcg	caggtgcagc	gcgggcgcct	ggggtcttgc	120
aaggctgagc	tgacgccgca	gaggtcgtgt	cacgtcccac	gaccttgacg	cogtcgggga	180
cagccggaac	agagcccggg	gaagcgggag	gcctcgggga	gcccctcggg	aagggcggcc	240
cgagagatac	gcaggtgcag	gtggccgcc				269

<210> 101

<211> 405

<212> DNA

<213> Homo sapien

<400> 101

tttttttttt	ttttggaatc	tactgcgagc	acagcaggtc	agcaacaagt	ttatttttgc	60
gctagcaagg	taacagggtg	gggcatgggt	acatgttcag	gtcaacttcc	tttgtogtgg	120
ttgattgggt	tgtctttatg	ggggcggggg	gggttagggg	aaacgaagca	aataacatgg	180

```
<210> 102
<211> 470
<212> DNA
<213> Homo sapien
```

```
<210> 103
<211> 581
<212> DNA
<213> Homo sapien
```

```
<210> 104
<211> 578
<212> DNA
<213> Homo sapien
```

<400> 104						
tttttttttt	tttttttttt	tttttctctt	cttttttttt	gaaatgagga	tcgagttttt	60
cactctctag	atagggcatg	aagaaaactc	atctttccag	ctttaaaata	acaatcaaat	120
ctcttatgct	atatcatatt	ttaagttaaa	ctaattgagtc	actggcttat	cttctcctga	180
aggaaatctg	ttcattcttc	tcattcatat	agttatatca	agtactacct	tgcatattga	240
gaggtttttc	ttctctattt	acacatatat	ttccatgtga	atttgatatca	aacctttatt	300
ttcatgcaaa	ctagaaaata	atgtttcttt	tgcataagag	aagagaacaa	tatagcatta	360
caaaactcgt	caaattgttt	gttaagttat	ccattataat	tagttggcag	gagctaatac	420
aaatcacatt	tacgacagca	ataataaaac	tgaagtacca	gttaaataatc	caaaataatt	480
aaaggaacat	ttttagcctg	ggtataatta	gctaattcac	tttacaagca	tttattagaa	540
tgaattcaca	tgttattatt	cttagcccaa	cacaatgg			578

<400> 105

```
<210> 106
<211> 473
<212> DNA
<213> Homo sapien
```

<400> 106

```
<210> 107
<211> 1621
<212> DNA
<213> Homo sapien
```

<400> 107

cgcctatggca	ctgcaggcca	tctcggtcat	ggagctgtcc	ggcctggccc	cgggcccggt	60
ctgtgctatg	gtcctggctg	acttcggggc	gcgtgtggta	cgcgtggacc	ggcccggtc	120
ccgctacgac	gtgagccgct	tgggcccggg	caagcgctcg	ctagtgtctg	acctgaagca	180
gccgcgggga	gccgcggtgc	tgcggcgctct	gtgcaagcgg	tccgatgtgc	tgtctggagcc	240
cttcgcgccg	ggtgtcatgg	agaaactcca	gctgggcccc	gagattctgc	agcgggaaaa	300
tccaaggcct	atttatgcca	ggctgagtg	atttggccag	tcaggaaagct	tctgccgggt	360
agctggccac	gatatcaact	at ttggcctt	gtcaggtgtt	ctctcaaaaa	ttggcagaag	420
tggtgagaat	ccgtatgccc	cgctgaatct	cctggctgac	tttgtgtgtg	gtggccttat	480
gtgtgcactg	ggcattataa	tggctctttt	tgaccgcaca	cgcactgaca	agggtcaggt	540
cattgatgca	aatatggtgg	aaggaacagc	atattttaagt	tcttttctgt	ggaaaactca	600
gaaatcgagt	ctgtgggaag	cacctcgagg	acagaacatg	ttggatggtg	gagcaccttt	660
ctatacgact	tacaggacag	cagatgggga	attcatggct	gttggaagca	tagaaccccc	720
gttctacgag	ctgctgatca	aaggacttgg	actaaagtct	gatgaacttc	ccaatcagat	780
gagcatggat	gattggccag	aaatgaagaa	gaagtttgca	gatgtatttg	caaagaagac	840
gaaggcgag	tgggtgtcaaa	tctttgacgg	cacagatgcc	tgtgtgaactc	cggttctgac	900
ttttgaggag	gttggttcac	atgatcaca	caaggaaacg	ggctcgttta	tcaccagtga	960
ggagcaggac	gtgagccccc	gccctgcacc	tctgtctgta	aacaccccag	ccatcccttc	1020

```

tttcaaaagg gatcctttca taggagaaca cactgaggag atacttgaag aatttggatt 1080
cagccgcgaa gagatttatc agcttaactc agataaaatc attgaaagta ataaggtaaa 1140
agctagtctc taacttccag gccacggct caagtgaatt tgaatactgc atttacagt 1200
tagagtaaca cataacattg tatgcatgga aacatggagg aacagtatta cagtgtccta 1260
ccactctaata caagaaaaga attacagact ctgattctac agtgatgatt gaattctaaa 1320
aatggttatc attagggctt ttgatttata aaactttggg tacttatact aaattatgg 1380
agttattctg ccttccagtt tgcttgatat atttggtgat attaagattc ttgacttata 1440
ttttgaatgg gttctagtga aaaaggaatg atatatctt gaagacatcg atatacattt 1500
atttacactc ttgattctac aatgtagaaa atgaggaaat gccacaaatt gtatggtgat 1560
aaaagtcacg tgaaacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1620
a 1621

```

```

<210> 108
<211> 382
<212> PRT
<213> Homo sapien

```

```

<400> 108
Met Ala Leu Gln Gly Ile Ser Val Met Glu Leu Ser Gly Leu Ala Pro
 1          5          10          15
Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
          20          25          30
Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
          35          40          45
Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
          50          55          60
Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
          65          70          75          80
Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
          85          90          95
Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
          100          105          110
Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
          115          120          125
Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
          130          135          140
Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met Cys
          145          150          155          160
Ala Leu Gly Ile Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Asp Lys
          165          170          175
Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu Ser
          180          185          190
Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro Arg
          195          200          205
Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr Arg
          210          215          220
Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln Phe
          225          230          235          240
Tyr Glu Leu Leu Ile Lys Gly Leu Gly Leu Lys Ser Asp Glu Leu Pro
          245          250          255
Asn Gln Met Ser Met Asp Asp Trp Pro Glu Met Lys Lys Lys Phe Ala
          260          265          270
Asp Val Phe Ala Lys Lys Thr Lys Ala Glu Trp Cys Gln Ile Phe Asp
          275          280          285

```

000000" 5'22'59.60"

Gly Thr Asp Ala Cys Val Thr Pro Val Leu Thr Phe Glu Glu Val Val
 290 295 300
 His His Asp His Asn Lys Glu Arg Gly Ser Phe Ile Thr Ser Glu Glu
 305 310 315 320
 Gln Asp Val Ser Pro Arg Pro Ala Pro Leu Leu Leu Asn Thr Pro Ala
 325 330 335
 Ile Pro Ser Phe Lys Arg Asp Pro Phe Ile Gly Glu His Thr Glu Glu
 340 345 350
 Ile Leu Glu Glu Phe Gly Phe Ser Arg Glu Glu Ile Tyr Gln Leu Asn
 355 360 365
 Ser Asp Lys Ile Ile Glu Ser Asn Lys Val Lys Ala Ser Leu
 370 375 380

<210> 109
 <211> 1524
 <212> DNA
 <213> Homo sapien

<400> 109
 ggcacgagggc tgcgccaggg cctgagcgga ggcggggggca gcctcgccag cggggggcccc 60
 gggcctggcc atgcctcact gagccagcgc ctgcgcctct acctcgccga cagctggaac 120
 cagtgcgacc tagtggctct cactgtcttc ctctggggcg tgggctgccg gctgaccccg 180
 ggtttgtagc acctggggcg cactgtcctc tgcctcgact tcatggtttt caccgtgctg 240
 ctgcttcaca tcttcacggc caacaaacag ctggggccca agatcgctcat cgtgagcaag 300
 atgatgaagg acgtgttctt ctctctcttc ttctctggcg tgtggctggt agcctatggc 360
 gtggccacgg aggggtcctt gaggccacgg gacagtgaact tcccaagtat cctgcgccgc 420
 gtcttctacc gtccctacct gcagatcttc gggcagattc cccaggagga catggacgtg 480
 gccctcatgg agcacagcaa ctgctcgtcg gagcccggtt tctgggcaca cctcctggg 540
 gcccaggcgg gcacctgcgt ctcccagtat gccaaactggc tgggtggtgct gctcctcgtc 600
 atcttctctg tctggcccaa catcctgctg gtcaacttgc tcattgccat gttcagttac 660
 acattcggca aagtacaggg caacagcgat ctctactgga aggcgcagcg ttaccgcctc 720
 atccgggaat tccactctcg gcccgcgctg gcccgcctt ttatcgctcat ctccacttg 780
 cgctctctgc tcaggcaatt gtgcaggcga ccccgagacc cccagccgtc ctccccggcc 840
 ctcgagcatt tccgggttta cctttctaag gaagccgagc ggaagctgct aacgtgggaa 900
 tccgtgcata aggagaactt tctgctggca cgcgctaggg acaagcggga gagcgactcc 960
 gagcgtctga agcgcacgtc ccagaaggtg gacttggcac tgaacagct gggacacatc 1020
 cgcgagtacg aacagcgctt gaaagtgtg gagcgggagg tccagcagt tagccgcgtc 1080
 ctgggggtggg tggccgaggg cctgagccgc tctgccttgc tgcccccagg tgggcccga 1140
 cccctgacc tgctgggtc caaagactga gccctgctgg cggacttcaa ggagaagccc 1200
 ccacagggga ttttgctcct agagtaaggc tcatctgggc ctcgggcccc gcacctggtg 1260
 gccttgctct tgaggtgagc cccatgtcca tctgggccac tgtcaggacc acctttggga 1320
 gtgtcatcct taaaaaccac agcatgcccg gctcctccca gaaccagtcc cagcctggga 1380
 ggatcaaggc ctggatcccg ggcggttatc catctggagg ctgcagggtc cttggggtaa 1440
 cagggaccac agaccctca ccactcacag attcctcaca ctggggaaat aaagccattt 1500
 cagagggaaa aaaaaaaaaa aaaa 1524

<210> 110
 <211> 3410
 <212> DNA
 <213> Homo sapien

<400> 110
 gggaaccagc ctgcacgcgc tggctccggg tgacagccgc gcgcctcggc caggatctga 60
 gtgatgagac gtgtccccac tgaggtgccc cacagcagca ggtgttgagc atgggctgag 120

agctggacc	ggcaccaaag	ggctggcaga	aatgggoccc	tggctgattc	ctaggcagtt	180
ggcggcagca	aggaggagag	gccgcagctt	ctggagcaga	gccgagacga	agcagttctg	240
gagtgcctga	acggcccccct	gagccctacc	cgcttgcccc	actatgggtcc	agaggctgtg	300
ggtgagccgc	ctgctgcggc	accggaaaagc	ccagctcttg	ctggtcaacc	tgctaacctt	360
tggcctggag	gtgtgttttg	ccgcaggcat	cacctatgtg	ccgcctctgc	tgctggaagt	420
gggggtagag	gagaagttca	tgaccatggt	gctgggcatt	ggtccagtg	tgggcctggt	480
ctgtgtccc	ctcctaggct	cagccagtg	ccactggcgt	ggacgctatg	gccgccgcg	540
gcccttcac	tgggcactgt	ccttgggcat	cctgctgagc	ctctttctca	tcccaagggc	600
cggctggcta	gcagggtctg	tgtgcccgga	tcccaggccc	ctggagctgg	cactgctcat	660
cctgggcgtg	gggctgctgg	acttctgtgg	ccaggtgtgc	ttcactccac	tggaggccct	720
gctctctgac	ctcttcggg	accggacca	ctgtcgccag	gcctactctg	tctatgcctt	780
catgatcagt	cttgggggct	gcctgggcta	cctcctgcct	gccattgact	gggacaccag	840
tgccctggcc	cctacctgg	gcacccagga	ggagtgcctc	tttggcctgc	tcacctcat	900
cttctcacc	tgcgtagcag	ccacactgct	ggtggctgag	gaggcagcgc	tgggccccac	960
cgagccagca	gaagggtgt	cggcccccctc	cttgtccccc	cactgtgtgc	catgcccggc	1020
ccgtctggct	tccggaacc	tgggcgcctc	gcttcgcccg	ctgcaccag	tgtgtctggc	1080
catgccccgc	accctgcgcc	ggctcttcgt	ggctgagctg	tgcagctgga	tggcactcat	1140
gaccttcacg	ctgttttaca	cggatttcgt	gggcgagggg	ctgtaccagg	gcgtgcccg	1200
agctgagccg	ggcaccgagg	cccgagaca	ctatgatgaa	ggcgttcgga	tgggcagcct	1260
ggggctgttc	ctgcagtgcg	ccatctccct	ggtcttctct	ctggtcatgg	accggctggt	1320
gcagcgattc	ggcactcgag	cagtctattt	ggccagtgtg	gcagctttcc	ctgtggctgc	1380
cgggtgccaca	tgctgtccc	acagtgtggc	cgtggtgaca	gcttcagccg	ccctcacccg	1440
gttcaccttc	tcagccctgc	agatcctgcc	ctacacactg	gcctccctct	accacccgga	1500
gaagcaggtg	ttcctgcca	aataccgagg	ggacactgga	ggtgctagca	gtgaggacag	1560
cctgatgacc	agcttcctgc	caggccctaa	gcctggagct	cccttcctca	atggacacgt	1620
gggtgctgga	ggcagtggcc	tgctcccacc	tccaccgcg	ctctgcgggg	cctctgcctg	1680
tgatgtctcc	gtacgtgtgg	tgggtgggtga	gcccaccgag	gccagggtgg	ttccggggccg	1740
gggcatctgc	ctggacctcg	ccatcctgga	tagtgecttc	ctgctgtccc	aggtggcccc	1800
atccctgttt	atgggctcca	ttgtccagct	cagccagtct	gtcactgcct	atatggtgtc	1860
tgccgcaggc	ctgggtctgg	tcgccattta	ctttgctaca	caggtagtat	ttgacaagag	1920
cgacttggcc	aaatactcag	cgtagaaaac	ttccagcaca	ttggggtgga	gggctgcct	1980
cactgggttc	cagctcccc	ctcctgttag	ccccatgggg	ctgcccggct	ggccgcagct	2040
ttctgttgct	gccaaagtaa	tgtggtctct	tgctgccacc	ctgtgctgct	gaggtgcgta	2100
cgtgcacag	tggggctgg	ggctgccttc	tctctctccc	ccagtctcta	gggtgcctg	2160
actgagggcc	ttccaagggg	gtttcagttc	ggacttatac	agggaggcca	gaagggtccc	2220
atgcactgga	atgcggggac	tctgcaggtg	gattaccag	gctcagggtt	aacagctagc	2280
ctcctagtty	agacacacct	agagaagggt	ttttgggagc	tgaataaact	cagtcacctg	2340
gtttcccatc	tctaagcccc	ttaacctgca	gcttcgttta	atgtagctct	tgcatgggag	2400
tttctaggat	gaaacactcc	tccatgggat	ttgaacatat	gacttatttg	taggggaaga	2460
gtcctgaggg	gcaacacaca	agaaccaggt	ccctcagcc	cacagcactg	tctttttgct	2520
gateccacccc	cctcttacct	tttatcagga	tgtggcctgt	tggtccttct	gttgccatca	2580
cagagacaca	ggcattttaa	tatttaactt	atttatttaa	caaagtagaa	gggaatccat	2640
tgctagcttt	tctgtgttgg	tgtctaatat	ttgggtaggg	tgggggatcc	ccaacaatca	2700
ggtccctcga	gatagctggt	cattgggctg	atcattgcca	gaatcttctt	ctcctggggg	2760
ctggccccc	aaaatgccta	accaggacc	ttggaaattc	tactcatccc	aatgataat	2820
tccaaatgct	gttacccaag	gttaggggtg	tgaaggaagg	tagagggtgg	ggcttcaggt	2880
ctcaacggct	tcctaacca	ccctctctct	cttggcccag	cctggttccc	cccacttcca	2940
ctccctctca	ctctctctag	gactgggctg	atgaaggcac	tgcccaaaat	ttcccttacc	3000
cccaactttc	cctaccccc	aactttcccc	accagctcca	caacctgtt	tggagctact	3060
gcaggaccag	aagcacaaag	tgcggtttcc	caagcctttg	tccatctcag	ccccagagt	3120
atatctgtgc	tggggaatc	tcacacagaa	actcaggagc	acccctgccc	tgagctaagg	3180
gaggtcttat	ctctcagggg	gggttttaagt	ccggttttga	ataatgtcgt	cttattttatt	3240
tagcgggggtg	aatattttat	actgtaagtg	agcaatcaga	gtataatggt	tatgggtgaca	3300
aaattaaaqq	ctttctttata	tqtttaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3360

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaataa aaaaaaaaaa

3410

<210> 111
<211> 1289
<212> DNA
<213> Homo sapien

<400> 111
agccaggcgt ccctctgcct gccactcag tggcaacacc cgggagctgt tttgtccttt 60
gtggagcctc agcagttccc tctttcagaa ctactgccca agagccctga acaggagcca 120
ccatgcagtg cttcagcttc attaagacca tgatgatcct cttcaatttg ctcactcttc 180
tgtgtggtgc agccctgttg gcagtgggca tctgggtgtc aatcgatggg gcactccttc 240
tgaagatctt cgggccactg tctgccagt ccagtcagtt tgtcaacgtg ggctacttcc 300
tcactgcagc cggcgttggt gtctttgtct ttggtttcct gggctgctat ggtgctaaga 360
ctgagagcaa gtgtgccctc gtgacgttct tcttcactct cctcctcatc ttcattgctg 420
aggttgcagc tgctgtggtc gccttgggtg acaccacaat ggctgagcac ttcctgacgt 480
tgctggtagt gcctgccatc aagaaagatt atggttccca ggaagacttc actcaagtgt 540
ggaacaccac catgaaaggg ctcaagtgtc gtggcttcac caactatacg gattttgagg 600
actcacccta cttcaaagag aacagtgcct ttccccatt ctgttgcaat gacaacgtca 660
ccaacacagc caatgaaacc tgcaccaagc aaaaggctca cgaccaaaaa gtagagggtt 720
gcttcaatca gcttttgtat gacatccgaa ctaatgcagt caccgtgggt ggtgtggcag 780
ctggaattgg gggcctcgag ctggctgccca tgattgtgtc catgtatctg tactgcaatc 840
tacaataagt ccacttctgc ctctgccact actgctgccca catgggaact gtgaagaggc 900
accctggcaa gcagcagtga ttgggggagg ggacaggatc taacaatgtc acttgggcca 960
gaatggacct gccctttctg ctccagactt ggggctagat agggaccact ccttttagcg 1020
atgctgactt ttccttccat tgggtgggtg atgggtgggg ggcattccag agcctctaag 1080
gtagccagtt ctgttgccca ttccccagt ctattaaacc cttgatatgc cccctaggcc 1140
tagtggtgat cccagtgtct tactggggga tgagagaaag gcattttata gcctgggcat 1200
aagtgaatc agcagagcct ctgggtggat gtgtagaagg cacttcaaaa tgcataaacc 1260
tgttacaatg ttaaaaaaaa aaaaaaaaaa 1289

<210> 112
<211> 315
<212> PRT
<213> Homo sapien

<400> 112
Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln
1 5 10 15
Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe
20 25 30
Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala
35 40 45
Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu
50 55 60
Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro
65 70 75 80
Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser
85 90 95
Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys
100 105 110
Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Val Ile Phe
115 120 125
Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe

```
<210> 113
<211> 553
<212> PRT
<213> Homo sapien
```

<400> 113															
Met	Val	Gln	Arg	Leu	Trp	Val	Ser	Arg	Leu	Leu	Arg	His	Arg	Lys	Ala
1				5					10					15	
Gln	Leu	Leu	Leu	Val	Asn	Leu	Leu	Thr	Phe	Gly	Leu	Glu	Val	Cys	Leu
			20					25					30		
Ala	Ala	Gly	Ile	Thr	Tyr	Val	Pro	Pro	Leu	Leu	Leu	Glu	Val	Gly	Val
		35					40					45			
Glu	Glu	Lys	Phe	Met	Thr	Met	Val	Leu	Gly	Ile	Gly	Pro	Val	Leu	Gly
	50					55					60				
Leu	Val	Cys	Val	Pro	Leu	Leu	Gly	Ser	Ala	Ser	Asp	His	Trp	Arg	Gly
65					70				75					80	
Arg	Tyr	Gly	Arg	Arg	Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu	Gly	Ile
			85				90						95		
Leu	Leu	Ser	Leu	Phe	Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala	Gly	Leu
		100					105					110			
Leu	Cys	Pro	Asp	Pro	Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile	Leu	Gly
		115					120					125			
Val	Gly	Leu	Leu	Asp	Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro	Leu	Glu
	130				135					140					
Ala	Leu	Leu	Ser	Asp	Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg	Gln	Ala
145					150					155				160	
Tyr	Ser	Val	Tyr	Ala	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu	Gly	Tyr
			165					170					175		
Leu	Leu	Pro	Ala	Ile	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro	Tyr	Leu
			180				185						190		

Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
 195 200 205
 Thr Cys Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly
 210 215 220
 Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
 225 230 235 240
 Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu
 245 250 255
 Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
 260 265 270
 Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
 275 280 285
 Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
 290 295 300
 Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
 305 310 315 320
 Val Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
 325 330 335
 Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg
 340 345 350
 Ala Val Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala
 355 360 365
 Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 370 375 380
 Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
 385 390 395 400
 Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
 405 410 415
 Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu
 420 425 430
 Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala
 435 440 445
 Gly Gly Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser
 450 455 460
 Ala Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
 465 470 475 480
 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 485 490 495
 Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
 500 505 510
 Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala
 515 520 525
 Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
 530 535 540
 Lys Ser Asp Leu Ala Lys Tyr Ser Ala
 545 550

<210> 114

<211> 241

<212> PRT

<213> Homo sapien

<400> 114

Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu

1 5 10 15
 Leu Ile Phe Leu Cys Gly Ala Ala Leu Leu Ala Val Gly Ile Trp Val
 20 25 30
 Ser Ile Asp Gly Ala Ser Phe Leu Lys Ile Phe Gly Pro Leu Ser Ser
 35 40 45
 Ser Ala Met Gln Phe Val Asn Val Gly Tyr Phe Leu Ile Ala Ala Gly
 50 55 60
 Val Val Val Phe Ala Leu Gly Phe Leu Gly Cys Tyr Gly Ala Lys Thr
 65 70 75 80
 Glu Ser Lys Cys Ala Leu Val Thr Phe Phe Phe Ile Leu Leu Leu Ile
 85 90 95
 Phe Ile Ala Glu Val Ala Ala Ala Val Val Ala Leu Val Tyr Thr Thr
 100 105 110
 Met Ala Glu His Phe Leu Thr Leu Leu Val Val Pro Ala Ile Lys Lys
 115 120 125
 Asp Tyr Gly Ser Gln Glu Asp Phe Thr Gln Val Trp Asn Thr Thr Met
 130 135 140
 Lys Gly Leu Lys Cys Cys Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp
 145 150 155 160
 Ser Pro Tyr Phe Lys Glu Asn Ser Ala Phe Pro Pro Phe Cys Cys Asn
 165 170 175
 Asp Asn Val Thr Asn Thr Ala Asn Glu Thr Cys Thr Lys Gln Lys Ala
 180 185 190
 His Asp Gln Lys Val Glu Gly Cys Phe Asn Gln Leu Leu Tyr Asp Ile
 195 200 205
 Arg Thr Asn Ala Val Thr Val Gly Gly Val Ala Ala Gly Ile Gly Gly
 210 215 220
 Leu Glu Leu Ala Ala Met Ile Val Ser Met Tyr Leu Tyr Cys Asn Leu
 225 230 235 240
 Gln

<210> 115
 <211> 366
 <212> DNA
 <213> Homo sapien

<400> 115
 gctctttctc tcccctcctc tgaatttaat tctttcaact tgcaatttgc aaggattaca 60
 catttcaactg tgatgtatat tgtgttgcaa aaaaaaaaaa gtgtctttgt ttaaaattac 120
 ttggtttggtg aatccatctt gctttttccc cattggaact agtcattaac ccactctctga 180
 actggtagaa aaacatctga agagctagtc tatcagcatc tgacagggtga attggatggt 240
 tctcagaacc atttcaccca gacagcctgt ttctatcctg tttaataaat tagtttggtg 300
 tctctacatg cataacaaac cctgctccaa tctgtcacat aaaagtctgt gacttgaagt 360
 ttagtc 366

<210> 116
 <211> 282
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(282)

009060"6225960

<223> n = A,T,C or G

<400> 116

```
acaaagatga accatttcct atattatagc aaaattaaaa tctacccgta ttctaataatt      60
gagaaatgag atnaaacaca atnttataaa gtctacttag agaagatcaa gtgacctcaa      120
agactttact attttcatat ttttaagacac atgattttatc ctatttttagt aacctgggtc      180
atacgttaaa caaaggataa tgtgaacagc agagaggatt tgttggcaga aaatctatgt      240
tcaatctnga actatctana tcacagacat ttctattcct tt                          282
```

<210> 117

<211> 305

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(305)

<223> n = A,T,C or G

<400> 117

```
acacatgtcg cttcactgcc ttcttagatg cttctgggtca acatanagga acagggacca      60
tatttatacct ccctcctgaa acaattgcaa aataanacaa aatatatgaa acaattgcaa      120
aataaggcaa aatatatgaa acaacaggtc tcgagatatt ggaaatcagt caatgaagga      180
tactgatccc tgatcactgt cctaatagcag gatgtgggaa acagatgagg tcacctctgt      240
gactgccccca gcttactgcc tgtagagagt ttctangctg cagttcagac agggagaaat      300
tgggt                                              305
```

<210> 118

<211> 71

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(71)

<223> n = A,T,C or G

<400> 118

```
accaaggtgt ntgaatctct gacgtgggga tctctgattc ccgcacaatc tgagtggaaa      60
aantcctggg t                                              71
```

<210> 119

<211> 212

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(212)

<223> n = A,T,C or G

<400> 119

```
actccggttg gtgtcagcag cacgtggcat tgaacatngc aatgtggagc ccaaaccaca      60
gaaaatgggg tgaaattggc caactttcta tnaacttatg ttggcaantt tgccaccaac      120
```

agtaagctgg cccttctaataaaaagaaaat tgaaagggtt ctcactaanc ggaattaant 180
aatggantca aganactccc aggcctcagc gt 212

<210> 120
<211> 90
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(90)
<223> n = A,T,C or G

<400> 120
actcgttgca natcaggggc cccccagagt caccgttgca ggagtccttc tggctcttgcc 60
ctccgccggc gcagaacatg ctgggggtgg 90

<210> 121
<211> 218
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(218)
<223> n = A,T,C or G

<400> 121
tgtancgtga anacgacaga nagggttgct aaaaatggag aanccttgaa gtcattttga 60
gaataagatt tgctaaaaga tttgggggcta aaacatgggt attgggagac atttctgaag 120
atatncangt aaattangga atgaattcat ggttcttttg ggaattcctt tacgatngcc 180
agcatanact tcatgtgggg atancagcta cccttgta 218

<210> 122
<211> 171
<212> DNA
<213> Homo sapien

<400> 122
taggggtgta tgcaactgta aggacaaaaa ttgagactca actggcttaa ccaataaagg 60
catttgtag tcatggaac aggaagtcgg atgggtgggc atcttcagt ctgcatgagt 120
caccaccccg gcgggggtcat ctgtgccaca ggtccctggt gacagtgcgg t 171

<210> 123
<211> 76
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(76)
<223> n = A,T,C or G

<400> 123

009060"6225960

```
<210> 124
<211> 131
<212> DNA
<213> Homo sapien
```

```
<210> 125
<211> 432
<212> DNA
<213> Homo sapien
```

```
<210> 126
<211> 112
<212> DNA
<213> Homo sapien
```

```
<210> 127
<211> 54
<212> DNA
<213> Homo sapien
```

```
<210> 128
<211> 323
<212> DNA
<213> Homo sapien
```

<400> 128						
acctcattag	taattgtttt	gttgtttcat	ttttttctaa	tgtctcccct	ctaccagctc	60
acctgagata	acagaatgaa	aatggaagga	cagccagatt	tctcctttgc	tctctgctca	120
ttctctctga	agtctaggtt	acccattttg	gggaccatt	ataggcaata	aacacagttc	180

```
<210> 129
<211> 192
<212> DNA
<213> Homo sapien
```

```

<400> 129
acatacatgt gtgtatatatt ttaaatatca cttttgtatc actctgactt tttagcatac      60
tgaaaacaca ctaacataat ttntgtgaac catgatcaga tacaacccaa atcattcatt      120
tagcacattc atctgtgata naaagatagg tgagtttcat ttccttcacg ttggccaatg      180
gataaacaaa      192

```

```
<220>  
<221> misc_feature  
<222> (1)...(362)  
<223> n = A,T,C or G
```

```
<210> 131
<211> 332
<212> DNA
<213> Homo sapien
```

<400> 131						
ctttttgaaa	gatcgtgtcc	actcctgtgg	acatcttgtt	ttaatggagt	ttcccatgca	60
gtangactgg	tatggttgca	gctgtccaga	taaaaacatt	tgaagagctc	caaaaatgaga	120
gtttctccag	gttcgcctcg	ctgcctcaag	tctcagcagc	agcctctttt	aggaggcatc	180
ttctgaacta	gattaagqca	qcttgtaaat	ctgatgtgat	ttgggtttatt	atccaactaa	240

<210> 135

```
<220>  
<221> misc_feature  
<222> (1)...(350)  
<223> n = A,T,C or G
```

```
<210> 136
<211> 399
<212> DNA
<213> Homo sapien
```

<400> 136						
tgtaccgtga	agacgacaga	agttgcatgg	cagggacagg	gcagggccga	ggccaggggt	60
gctgtgattg	tatccgaata	ntcctcgtga	gaaaagataa	tgagatgacg	tgagcagcct	120
gcagacttgt	gtctgccttc	aanaagccag	acaggaaggc	cctgcctgcc	ttggctctga	180
cctggcggcc	agccagccag	ccacaggttg	gcttcttcct	tttgtggtga	caacnccaag	240
aaaactgcag	agggcccagg	tcaggtgtga	gtgggtangt	gaccataaaa	caccaggtgc	300
tccaggaac	ccgggcaaag	gccatcccca	cctacagcca	gcatgccccac	tggcgtgatg	360
ggtgcagang	gatgaagcag	ccagntgttc	tgctgtggt			399

```
<220>  
<221> misc_feature  
<222> (1)...(165)  
<223> n = A,T,C or G
```

<210> 138
<211> 338
<212> DNA

<223> n = A, T, C or G

actcactgga	atgccacatt	cacaacagaa	tcagaggtct	gtgaaaacat	taatggctcc	60
ttaacttctc	cagtaagaat	cagggacttg	aaatggaaac	gttaacagcc	acatgccccaa	120
tgctgggcag	tctcccatgc	cttcacagt	gaaagggctt	gagaaaaatc	acatccaatg	180
tcatgtgttt	ccagccacac	caaaaggtgc	ttgggggtga	gggctggggg	catananggt	240
cangctcag	gaagcctcaa	gttcattca	gctttgccac	tgtacattcc	ccatntttta	300
aaaaactgat	gccttttttt	tttttttttg	taaaattc			338

<213> Homo sapien

gggaatcttg	gtttttggca	tctggtttgc	ctatagccga	ggccactttg	acagaacaaa	60
gaaagggact	tcgagtaaga	aggtgattta	cagccagcct	agtgcccgaa	gtgaaggaga	120
attcaaacag	acctcgtcac	tcctgggttg	agcctggctg	gctcacgcgc	tatcatctgc	180
atttgcccta	ctcaggtgct	accggactct	ggccccgat	gtctgtagtt	tcacaggatg	240
ccttatttgt	cttctacacc	ccacagggcc	ccctacttct	tcggatgtgt	ttttaataat	300
gtcagctatg	tgccccatcc	tccttcatgc	cctccctccc	tttccctacca	ctgctgagtg	360
gcctggaact	tgtttaaaagt	gt				382

<213> Homo sapien

<223> n = A, T, C or G

acccaaanctt	ctttctgttg	tgttngattt	tactataggg	gtttngcttn	ttctaaanat	60
acttttcatt	taacancttt	tgттаagtgt	caggctgcac	tttgctccat	anaattattg	120
ttttcacatt	tcaacttgta	tgtgtttgtc	tcttanagca	ttggtgaaat	cacatatattt	180
atattcagca	taaaggagaa					200

<213> Homo sapien

<223> n = A, T, C or G

<400> 141

<211> 459

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

<400> 142

<211> 140

<213> Homo sapien

<400> 143

<211> 164

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

<400> 144

<210> 145
 <211> 303
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(303)
 <223> n = A,T,C or G

<400> 145
 acgtagacca tccaactttg tatttgtaat ggcaaacatc cagnagcaat toctaaacaa 60
 actggagggt atttataccc aattatccca ttcattaaca tgccctcctc ctcaggctat 120
 gcaggacagc tatcataagt cggcccaggc atccagatac taccatttgt ataaacttca 180
 gtaggggagt ccatccaagt gacaggtcta atcaaaggag gaaatggaac ataagcccag 240
 tagtaaaatn ttgcttagct gaaacagcca caaaagactt accgccgtgg tgattaccat 300
 caa 303

<210> 146
 <211> 327
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(327)
 <223> n = A,T,C or G

<400> 146
 actgcagctc aattagaagt ggtctctgac tttcatcanc ttctccctgg gtcctatgac 60
 actggcctgg agtgactcat tgctctgggt gggtgagaga gtccttttgc caacaggcct 120
 ccaagtcagg gctgggattt gtttcccttc cacattctag caacaatatg ctggccactt 180
 cctgaacagg gaggggtggga ggagccagca tggaacaagc tgccactttc taaagtagcc 240
 agacttgccc ctgggcctgt cacacctact gatgaccttc tgtgcctgca ggatggaatg 300
 taggggtgag ctgtgtgact ctatggt 327

<210> 147
 <211> 173
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(173)
 <223> n = A,T,C or G

<400> 147
 acattgtttt tttgagataa agcattgana gagctctcct taacgtgaca caatggaagg 60
 actggaacac ataccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120
 atattcaagc acatatgtta tatattatc agttccatgt ttatagccta gtt 173

<210> 148
 <211> 477
 <212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(477)

<223> n = A,T,C or G

<400> 148

acaaccactt	tatctcatcg	aatttttaac	ccaaactcac	tactgtgcc	tttctatcct	60
atgggatata	ttatttgatg	ctccatttca	tcacacatat	atgaataata	cactcatact	120
gccctactac	ctgctgcaat	aatcacattc	ccttcctgtc	ctgaccctga	agccattggg	180
gtggtcctag	tggccatcag	tccangcctg	caccttgagc	ccttgagctc	cattgctcac	240
nccancccac	ctcaccgacc	ccatcctctt	acacagctac	ctccttgctc	tctaaccoca	300
tagattatnt	ccaaattcag	tcaattaagt	tactattaac	actctaccgc	acatgtccag	360
caccactggg	aagccttctc	cagccaacac	acacacacac	acacncacac	acacacatat	420
ccaggcacag	gctacctcat	cttcacaatc	acccttttaa	ttaccatgct	atggtgg	477

<210> 149

<211> 207

<212> DNA

<213> Homo sapien

<400> 149

acagttgtat	tataatatca	agaaataaac	ttgcaatgag	agcattttaag	agggagaac	60
taacgtat	tagagagcca	aggaagggtt	ctgtggggag	tgggatgtaa	ggtagggcct	120
gatgataaat	aagagtcagc	caggtaagt	ggtggtgtgg	tatgggcaca	gtgaagaaca	180
tttcaggcag	agggaacagc	agtgaaa				207

<210> 150

<211> 111

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(111)

<223> n = A,T,C or G

<400> 150

accttgattt	cattgctgct	ctgatggaaa	cccaactatc	taatttagct	aaaacatggg	60
cacttaa	atg	tggtcagtgt	ttggacttgt	taactantgg	catctttggg	111

<210> 151

<211> 196

<212> DNA

<213> Homo sapien

<400> 151

agcgcgag	gtcatattga	acattccaga	tacctatcat	tactcgatgc	tgttgataac	60
agcaagatgg	ctttgaactc	agggtcacca	ccagctattg	gaccttacta	tgaaaaccat	120
ggataccaac	cggaaaaccc	ctatcccgca	cagccactg	tgggtccccc	tgtctacgag	180
gtgcatccgg	ctcagt					196

<210> 152

actggaaata	ataaaaccca	catcacagtg	ttgtgtcaaa	gatcatcagg	gcatggatgg	60
gaaagtgctt	tgggaactgt	aaagtgccta	acacatgata	gatgattttt	gttataatat	120
ttgaatcacg	gtgcatacaa	actctcctgc	ctgctcctcc	tgggccccag	ccccagcccc	180
atcacagctc	actgctctgt	tcataccaggc	ccagcatgta	gtggetgatt	cttcttggct	240

```
<220>
<221> misc_feature
<222> (1)...(498)
<223> n = A,T,C or G
```



```

<400> 163
catttataca gacaggcgtg aagacattca cgacaaaaac gcgaaattct atcccgtagc      60
canagaaggc agctacggct actcctacat cctggcgtgg gtggccttcg cctgcacctt    120
catcagcggc atgatgt                                     137

```

```

<220>
<221> misc_feature
<222> (1)...(469)
<223> n = A,T,C or G

<400> 164
cttatcacaa tgaatgttct cctgggcagc gttgtgatct ttgccacctt cgtgacttta      60
tgcaatgcat catgctatatt catacctaata gagggagttc caggagattc aaccaggaaa      120
tgcatggatc tcaaaggaaa caaacaccca ataaactcgg agtggcagac tgacaactgt      180
gagacatgca cttgctacga aacagaaatt tcatgttgca cccttgtttc tacacctgtg      240
ggttatgaca aagacaactg ccaaagaatc ttcaagaagg aggactgcaa gtatatcgtg      300
gtggagaaga aggagcccaa aaagacctgt tctgtcagtg aatggataat ctaatgtgct      360
tctagtaggc acagggctcc caggccaggc ctcattctcc tctggcctct aatagtcaat      420
gatttgttag ccatgcctat cagtaaaaag atntttgagc aaacacttt      469

```

```

<220>
<221> misc_feature
<222> (1)...(195)
<223> n = A,T,C or G

<400> 165
acagtttttt atatatatcg acattgccgg cacttggtgtt cagtttcata aagctgggtgg      60
atccgctgtc atccactatt ccttggttag agtaaaaatt attcttatag cccatgtccc      120
tgcaggccgc ccgcccgtag ttctcgttcc agtcgtcttg gcacacaggg tgccaggact      180
tcctctqaga tgagt                                     195

```

```
<220>
<221> misc feature
```

<222> (1)...(383)

<223> n = A,T,C or G

<400> 166

acatcttagt	agtgtggcac	atcagggggc	catcaggggc	acagtcactc	atagcctcgc	60
cgaggtcggg	gtccacacca	ccggtgtagg	tgtgctcaat	cttgggcttg	gcgcccacct	120
ttggagaagg	gatatgctgc	acacacatgt	ccacaaagcc	tgtgaactcg	ccaaagaatt	180
tttgagacc	agcctgagca	aggggcggat	gttcagcttc	agctcctcct	tcgtcagggtg	240
gatgcccaacc	tcgtctangg	tccgtgggaa	gctgggtgtcc	acntcaccta	caacctgggc	300
gangatctta	taaagaggct	ccnagataaa	ctccacgaaa	cttctctggg	agctgctagt	360
nggggccttt	ttggtgaact	ttc				383

<210> 167

<211> 247

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(247)

<223> n = A,T,C or G

<400> 167

acagagccag	accttggcca	taaatgaanc	agagattaag	actaaacccc	aagtcganat	60
tggagcagaa	actggagcaa	gaagtgggcc	tggggctgaa	gtagagacca	aggccactgc	120
tatanccata	cacagagcca	actctcaggc	caaggcnatg	gttggggcag	anccagagac	180
tcaatctgan	tccaaagtgg	tggctggaac	actggtcatg	acanaggcag	tgactctgac	240
tgangtc						247

<210> 168

<211> 273

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(273)

<223> n = A,T,C or G

<400> 168

acttctaagt	tttctagaag	tggaaggatt	gtantcatcc	tgaaaatggg	tttacttcaa	60
aatccctcan	ccttggttctt	cacnactgtc	tatactgana	gtgtcatggt	tcacaaaagg	120
gctgacacct	gagcctgnat	tttcaactcat	ccctgagaag	ccctttccag	taggggtgggc	180
aattcccaac	ttccttgcca	caagcttccc	aggctttctc	ccctggaaaa	ctccagcttg	240
agtcccagat	acactcatgg	gctgccctgg	gca			273

<210> 169

<211> 431

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(431)

<223> n = A,T,C or G

<400> 169

acagccttgg	cttccccaaa	ctccacagtc	tcagtgcaga	aagatcatct	tccagcagtc	60
agctcagacc	aggggtcaaag	gatgtgacat	caacagtttc	tggtttcaga	acaggttcta	120
ctactgtcaa	atgaccccc	atacttcctc	aaaggctgtg	gtaagttttg	cacaggtgag	180
ggcagcagaa	aggggggtant	tactgatgga	caccatcttc	tctgtatact	ccacactgac	240
cttgccatgg	gcaaaggccc	ctaccacaaa	aacaatagga	tcactgctgg	gcaccagctc	300
acgcacatca	ctgacaaccg	ggatggaaaa	agaantgcc	actttcatac	atccaactgg	360
aaagtgatct	gatactggat	tcttaattac	cttcaaaagc	ttctgggggc	catcagctgc	420
tcgaacactg	a					431

<210> 170

<211> 266

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(266)

<223> n = A,T,C or G

<400> 170

acctgtgggc	tgggctgtta	tgectgtgcc	ggctgctgaa	agggagttca	gaggtggagc	60
tcaaggagct	ctgcaggcat	tttgccaanc	ctctccanag	canagggagc	aacctacact	120
ccccgctaga	aagacaccag	attggagtcc	tgggaggggg	agttgggggtg	ggcatttgat	180
gtatacttgt	cacctgaatg	aangagccag	agaggaanga	gacgaanatg	anattggcct	240
tcaaagctag	gggtctggca	ggtgga				266

<210> 171

<211> 1248

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(1248)

<223> n = A,T,C or G

<400> 171

ggcagccaaa	tcataaacgg	cgaggactgc	agcccgcact	cgcagccctg	gcaggcggca	60
ctggtcatgg	aaaacgaatt	gttctgctcg	ggcgtcctgg	tgcacccgca	gtgggtgctg	120
tcagccgcac	actgtttcca	gaagtgagtg	cagagctcct	acaccatcgg	gctgggcctg	180
cacagtcttg	aggccgacca	agagccaggg	agccagatgg	tggaggccag	cctctccgta	240
cggcaccag	agtacaacag	acccttgctc	gctaacgacc	tcattgctcat	caagttggac	300
gaatccgtgt	ccgagtctga	caccatccgg	agcatcagca	ttgcttcgca	gtgccctacc	360
gcggggaact	cttgccctgt	ttctggctgg	ggtctgctgg	cgaacggcag	aatgcctacc	420
gtgctgcagt	gcgtgaacgt	gtcgggtggtg	tctgaggagg	tctgcagtaa	gctctatgac	480
ccgtgttacc	accccagcat	gttctgcgcc	ggcggagggc	aagaccagaa	ggactcctgc	540
aacggtgact	ctggggggcc	cctgatctgc	aacgggtact	tgcagggcct	tgtgtctttc	600
ggaaaagccc	cgtgtggcca	agttggcgtg	ccagggtgtct	acaccaacct	ctgcaaattc	660
actgagtgga	tagagaaaac	cgtccaggcc	agttaactct	ggggactggg	aacccatgaa	720
attgaccccc	aaatacatcc	tgcggaagga	attcaggaat	atctgttccc	agccccctct	780
ccctcaggcc	caggagtcca	ggcccccagc	ccctcctccc	tcaaaccaag	ggtacagatc	840


```

cccagcccct cctccctcag acccaggagt ccagaccccc cagcccctcc tccctcagac 900
ccaggagtcc agcccctcct ccctcagacc caggagtcca gacccccccag cccctcctcc 960
ctcagaccca ggggtccagg cccccaaccc ctcctccctc agactcagag gtccaagccc 1020
ccaaccntc attccccaga cccagagggtc cagggtcccag cccctcntcc ctcagaccca 1080
gcggtccaat gccacctaga ctntccctgt acacagtgcc cccttgtggc acgttgaccc 1140
aaccttacca gttggttttt catttttngt ccctttcccc tagatccaga aataaagttt 1200
aagagaagng caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1248

```

<210> 172

<211> 159

<212> PRT

<213> Homo sapien

<220>

<221> VARIANT

<222> (1)...(159)

<223> Xaa = Any Amino Acid

<400> 172

```

Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro
 1          5          10          15
Leu Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser
          20          25          30
Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr
          35          40          45
Ala Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly
          50          55          60
Arg Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu
          65          70          75          80
Glu Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe
          85          90          95
Cys Ala Gly Gly Gly Gln Xaa Gln Xaa Asp Ser Cys Asn Gly Asp Ser
          100         105         110
Gly Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe
          115         120         125
Gly Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn
          130         135         140
Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
          145         150         155

```

<210> 173

<211> 1265

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(1265)

<223> n = A,T,C or G

<400> 173

```

ggcagcccgc actgcgagcc ctggcaggcg gcactgggtca tggaaaacga attgttctgc 60
tggggcgctc tgggtgcatcc gcagtgggtg ctgtcagccg cacactgttt ccagaactcc 120
tacaccatcg ggctgggcct gcacagtctt gaggccgacc aagagccagg gagccagatg 180

```

```
<210> 174
<211> 1459
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1459)
<223> n = A,T,C or G
```

<400> 174						
ggtcagccgc	acactgtttc	cagaagtgag	tgcagagctc	ctacaccatc	gggctgggcc	60
tgcacagtct	tgaggccgac	caagagccag	ggagccagat	ggtggaggcc	agcctctccg	120
tacggcacc	agagtacaac	agacccttgc	tcgctaacga	cctcatgctc	atcaagttgg	180
acgaatccgt	gtccgagtct	gacaccatcc	ggagcatcag	cattgcttcg	cagtgcctta	240
ccgcggggaa	ctcttgccct	gtttctggct	ggggtctgct	ggcgaacggg	gagctcacgg	300
gtgtgtgtct	gccctcttca	aggaggtcct	ctgcccagtc	gcgggggctg	accagagct	360
ctgcgtcca	ggcagaatgc	ctaccgtgct	gcagtgcgtg	aacgtgtcgg	tgggtgtctga	420
ngaggtctgc	antaagctct	atgacccgct	gtaccacccc	ancatgttct	gcgcggcg	480
agggcaagac	cagaaggact	cctgcaacgt	gagagagggg	aaaggggagg	gcaggcgact	540
cagggaaggg	tggagaaggg	ggagacagag	acacacaggg	ccgcatggcg	agatgcagag	600
atggagagac	acacaggggag	acagtgacaa	ctagagagag	aaactgagag	aaacagagaa	660
ataaacacag	gaataaagag	aagcaaagga	agagagaaac	agaaacagac	atggggaggc	720
agaaacacac	acacatagaa	atgcagttga	ccttccaaca	gcatggggcc	tgagggcggt	780
gacctccacc	caatagaaaa	tctctttata	acttttgact	ccccaaaaac	ctgactagaa	840
atagcctact	gttgacgggg	agcctttacca	ataacataaa	tagtcgattt	atgcatacgt	900
tttatgcat	catgatatac	ctttgttgga	attttttgat	atttctaagc	tacacagttc	960
gtctgtgaat	ttttttaaat	tgttgcacat	ctcctaaaat	tttctgtag	tggtttattga	1020
aaaaatccaa	gtataagtgg	acttggcat	tcaaaccagg	gttgttcaag	ggtcaactgt	1080
taccagag	ggaaacagtg	acacagattc	atagaggtga	aacacgaaga	gaaacaggaa	1140
aaatcaagac	tctacaaaga	ggctgggcag	ggtggctcat	gcctgtaatc	ccagcacttt	1200
gggaggcgag	gcaggcgagat	cacttgaggt	aaggagttca	agaccagcct	ggccaaaatg	1260
gtgaaatcct	gtctgtacta	aaaatacaaa	agttagctgg	atatggtggc	aggcgctgt	1320
aatcccagct	acttgggagg	ctgaggcgag	agaattgctt	gaatatggga	ggcagaggtt	1380

gaagtgagtt gagatcacac cactatactc cagctggggc aacagagtaa gactctgtct 1440
caaaaaaaaa aaaaaaaaaa 1459

<210> 175
<211> 1167
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1167)
<223> n = A,T,C or G

<400> 175
ggcagccct ggcaggcggc actggtcatg gaaaacgaat tgttctgctc gggcgtcctg 60
gtgcatccgc agtgggtgct gtcagccgca cactgtttcc agaactccta caccatcggg 120
ctgggcctgc acagtcttga ggccgaccaa gagccaggga gccagatggt ggaggccagc 180
ctctccgtac ggcaccaga gtacaacaga ctcttgctcg ctaacgacct catgctcatc 240
aagttggacg aatccgtgtc cgagtctgac accatccgga gcatcagcat tgcttcgcag 300
tgccctaccg cggggaactc ttgcctcgtn tctggctggg gtctgctggc gaacggcaga 360
atgcctaccg tgctgcaactg cgtgaacgtg tccgtgggtg ctgaggangt ctgcagtaag 420
ctctatgacc cgctgtacca cccagcatg ttctgcgccg gcggaggggca agaccagaag 480
gactcctgca acggtgactc tggggggccc ctgatctgca acgggtactt gcagggcctt 540
gtgtctttcg gaaaagcccc gtgtggccaa cttggcgtgc cagggtgtcta caccaacctc 600
tgcaaattca ctgagtggat agagaaaacc gtccagncca gttaactctg gggactggga 660
acccatgaaa ttgaccccca aatacatcct gcggaangaa ttcaggaata tctgttccca 720
gccccctcct cctcaggccc aggagtcag gccccagcc cctcctccct caaaccaagg 780
gtacagatcc ccagccctc ctccctcaga ccaggagtc cagaccccc agccctcnt 840
cctcagacc caggagtcca gccccctcct cntcagagc aggagtcag accccccagc 900
cctcntccg tcagaccag ggggtgcagg ccccaacccc tcntcctca gagtcagagg 960
tccaagcccc caaccctcg ttccccagac ccagaggtnc aggtcccag cctcctccc 1020
tcagaccag cgggtccaat ccacctagan tntccctgta cacagtgcc ccttgtggca 1080
ngttgacca accttaccag ttggtttttc attttttgtc cctttccct agatccagaa 1140
ataaagtnta agagaagcgc aaaaaaa 1167

<210> 176
<211> 205
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (1)...(205)
<223> Xaa = Any Amino Acid

<400> 176
Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp
1 5 10 15
Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu
20 25 30
Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val
35 40 45
Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Leu Leu Leu
50 55 60

Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
 65 70 75 80
 Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
 85 90 95
 Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg Met
 100 105 110
 Pro Thr Val Leu His Cys Val Asn Val Ser Val Val Ser Glu Xaa Val
 115 120 125
 Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys Ala
 130 135 140
 Gly Gly Gly Gln Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly Gly
 145 150 155 160
 Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly Lys
 165 170 175
 Ala Pro Cys Gly Gln Leu Gly Val Pro Gly Val Tyr Thr Asn Leu Cys
 180 185 190
 Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Xaa Ser
 195 200 205

<210> 177
 <211> 1119
 <212> DNA
 <213> Homo sapien

<400> 177
 gcgcactcgc agccctggca ggcggcactg gtcattgaaa acgaattggt ctgctcgggc 60
 gtccctgggtgc atccgcagtg ggtgctgtca gccgcacact gtttccagaa ctccacaccc 120
 atcgggctgg gcctgcacag tcttgaggcc gaccaagagc cagggagcca gatggtggag 180
 gccagcctct ccgtacggca cccagagtac aacagaccct tgctcgctaa cgacctcatg 240
 ctcactcaagt tggacgaatc cgtgtccgag tctgacacca tccggagcat cagcattgct 300
 tcgcagtgcc ctaccgctgg gaactcttgc ctcgtttctg gctggggtct gctggcgaac 360
 gatgctgtga ttgccatcca gtcccagact gtgggaggct gggagtgtga gaagctttcc 420
 caaccctggc aggggtgtac catttcggca acttccagtg caaggacgtc ctgctgcac 480
 ctactgggt gctcactact gctcactgca tcaccggaa cactgtgatc aactagccag 540
 caccatagtt ctccgaagtc agactatcat gattactgtg ttgactgtgc tgtctattgt 600
 actaaccatg ccgatgttta ggtgaaatta gcgtcacttg gcctcaacca tcttggtatc 660
 cagttatcct cactgaattg agatttcctg cttcagtgtc agccattccc acataatttc 720
 tgacctacag aggtgagga tcatatagct cttcaaggat gctggtactc ccctcacaaa 780
 ttcattttctc ctgttgtagt gaaagggtgc ccctctggag cctcccaggg tgggtgtgca 840
 ggtcacaatg atgaatgtat gatcgtgttc ccattaccca aagccttta atccctcatg 900
 ctcagtacac cagggcaggt ctagcatttc ttcatttagt gtatgctgtc cattcatgca 960
 accacctcag gactcctgga ttctctgcct agttgagctc ctgcatgctg cctccttggg 1020
 gaggtgaggg agagggccca tggttcaatg ggatctgtgc agttgtaaca cattaggtgc 1080
 ttaataaaca gaagctgtga tgttaaaaaa aaaaaaaaaa 1119

<210> 178
 <211> 164
 <212> PRT
 <213> Homo sapien

<220>
 <221> VARIANT
 <222> (1) ... (164)
 <223> Xaa = Any Amino Acid

000000"025960

<400> 178
Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp
1 5 10 15
Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu
20 25 30
Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val
35 40 45
Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu Leu
50 55 60
Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
65 70 75 80
Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
85 90 95
Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Asp Ala Val
100 105 110
Ile Ala Ile Gln Ser Xaa Thr Val Gly Gly Trp Glu Cys Glu Lys Leu
115 120 125
Ser Gln Pro Trp Gln Gly Cys Thr Ile Ser Ala Thr Ser Ser Ala Arg
130 135 140
Thr Ser Cys Cys Ile Leu Thr Gly Cys Ser Leu Leu Leu Thr Ala Ser
145 150 155 160
Pro Gly Thr Leu

<210> 179
<211> 250
<212> DNA
<213> Homo sapien

<400> 179
ctggagtgcc ttggtgtttc aagcccctgc aggaagcaga atgcaccttc tgaggcacct 60
ccagctgccc ccggccgggg gatgagaggc tcggagcacc cttgcccggc tgtgattgct 120
gccaggcact gttcatctca gcttttctgt ccctttgctc ccggcaagcg cttctgctga 180
aagttcatat ctggagcctg atgtcttaac gaataaaggt cccatgctcc acccgaaaaa 240
aaaaaaaaa 250

<210> 180
<211> 202
<212> DNA
<213> Homo sapien

<400> 180
actagtccag tgtggtggaa ttccattgtg ttgggcccac cacaatggct acctttaaca 60
tcacccagac cccgcccctg cccgtgcccc acgtgtgtgc taacgacagt atgatgttta 120
ctctgtact cggaactat ttttatgtaa ttaatgtatg ctttcttgtt tataaatgcc 180
tgattttaaaa aaaaaaaaaa aa 202

<210> 181
<211> 558
<212> DNA
<213> Homo sapien

<220>

000000" 5225960

<400> 181

<210> 182

<211> 479

<212> DNA

<213> Homo sapien

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (479)$

<223> n = A, T, C or G

<400> 182

acagggwttk	grggatgcta	agscgccrga	rwtygtttga	tccaaccctg	gcttwttttc	60
agaggggaaa	atggggccta	gaagttacag	mscatytagy	tggtgcgmtg	gcacccctgg	120
cstcacacag	astcccgagt	agctgggact	acaggcacac	agtactgaa	gcaggccctg	180
ttwgcaattc	acgttgccac	ctccaactta	aacattcttc	atatgtgatg	tccttagtca	240
ctaaggttaa	actttcccac	ccagaaaagg	caacttagat	aaaatcttag	agtactttca	300
tactmttcta	agtctctctc	cagcctcact	kgagtcctm	cytgggggtt	gataggaant	360
ntctcttggc	tttctcaata	aartctctat	ycatctcatg	tttaatttgg	tacgcatara	420
awtgstgara	aaattaaaaa	gttctggtty	mactttaaaa	aaaaaaaaaa	aaaaaaaaaa	479

<210> 183

<211> 384

<212> DNA

<213> Homo sapien

<400> 183

aggcgggagc	agaagctaaa	gccaaagccc	aagaagagtg	gcagtgccag	cactggtgcc	60
agtaccagta	ccaataacag	tgccagtgcc	agtgccagca	ccagtgggtg	cttcagtgtc	120
ggtgccagcc	tgaccgccac	tctcacattt	gggctcttcg	ctggccttgg	tggagctggt	180
gccagcacca	gtggcagctc	tgggtgctgt	ggtttctcct	acaagtgaga	ttttagatat	240
tgtaaactct	gccagtcttt	ctcttcaagc	cagggtgcat	cctcagaaac	ctactcaaca	300
cagcactcta	ggcagccact	atcaatcaat	tgaagttgac	actctgcatt	aratctattt	360
gccatttcaa	aaaaaaaaaa	aaaa				384

<210> 184

<211> 496

<212> DNA

<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(496)
 <223> n = A,T,C or G

<400> 184
 accgaattgg gaccgctggc ttataagcga tcatgttynt ccrgtatcac ctcaacgagc 60
 agggagatcg agtctatacg ctgaagaaat ttgacccgat gggacaacag acctgctcag 120
 cccatcctgc tcggtttctcc ccagatgaca aatactctsg acaccgaatc accatcaaga 180
 aacgcttcaa ggtgctcatg acccagcaac cgcgccctgt cctctgaggg tcccttaaac 240
 tgatgtcttt tctgccacct gttaccacct ggagactccg taaccaaact cttcggactg 300
 tgagccctga tgcttttttg ccagccatac tctttggcat ccagtctctc gtggcgattg 360
 attatgcttg tgtgaggcaa tcatggtggc atcacccata aagggaacac atttgacttt 420
 tttttctcat attttaaatt actacmagaw tattwmagaw waaatgawtt gaaaaactst 480
 taaaaaaaaa aaaaaa 496

<210> 185
 <211> 384
 <212> DNA
 <213> Homo sapien

<400> 185
 gctggtagcc tatggcgkgg cccacggagg ggctcctgag gccacggrac agtgacttcc 60
 caagtatcyt ggcsgcgtc ttctaccgtc cctacctgca gatcttcggg cagattcccc 120
 agggagacat ggacgtggcc ctcatggagc acagcaactg ytcgtcggag ccgggtttct 180
 gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc aactggctgg 240
 tgggtgctgct cctcgtcatc ttctgctcgt tggccaacat cctgctgggc aacttgetca 300
 ttgccatgtt cagttacaca ttcggcaaag tacagggcaa cagcgatctc tactgggaag 360
 ggcgcagcgtt accgcctcat ccgg 384

<210> 186
 <211> 577
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(577)
 <223> n = A,T,C or G

<400> 186
 gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc 60
 tnccatcgtc atactgtagg tttgccacca cytcctggca tcttggggcg gcntaatatt 120
 ccaggaaact ctcaatcaag tcaccgtcga tgaaacctgt gggctggttc tgtcttcgcg 180
 tcggtgtgaa aggatctccc agaaggagtg ctcgatcttc cccacacttt tgatgacttt 240
 attgagtcca ttctgcatgt ccagcaggag gttgtaccag ctctctgaca gtgaggtcac 300
 cagccctatc atgccgttga mcgtgccgaa garcaccgag ccttgtgtgg gggkkgaagt 360
 ctcaccacaga ttctgcatta ccagagagcc gtggcaaaag acattgacaa actcgcacag 420
 gtggaaaaag amcamctcct ggargtgctn gccgctcctc gtcmttggt ggcagcgctw 480
 tccttttgac acacaaacaa gttaaaggca ttttcagccc ccagaaantt gtcacatcc 540
 aagatntcgc acagcactna tccagttggg attaaat 577

<210> 187

```
<220>
<221> misc_feature
<222> (1)...(534)
<223> n = A,T,C or G
```

```
<210> 188
<211> 761
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(761)
<223> n = A,T,C or G
```

```
<210> 189
<211> 482
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G
```


<400> 189
 tttttttttt tttgccgatn ctactatttt attgcaggan gtgggggtgt atgcaccgca 60
 caccggggct atnagaagca agaaggaagg agggagggca cagccccttg ctgagcaaca 120
 aagccgcctg ctgccttctc tgtctgtctc ctggtgcagg cacatgggga gaccttcccc 180
 aaggcagggg ccaccagtcc aggggtggga atacaggggg tgggagtgt gcataagaag 240
 tgataggcac agggcaccgg gtacagaccc ctcggtcctt gacaggtnga ttctgaccag 300
 gtcattgtgc cctgccagg cacagcgtan atctggaaaa gacagaatgc ttcccttttc 360
 aaatttggtc ngtcatngaa ngggcanttt tccaanttng gctnggtctt ggtacncttg 420
 gttcggccca gtcncnctc caaaaantat tcaccnct ccnaattgct tgcnggnccc 480
 cc 482

<210> 190
 <211> 471
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(471)
 <223> n = A,T,C or G

<400> 190
 tttttttttt ttttaaaaca gtttttcaca acaaaattta ttagaagaat agtgggtttg 60
 aaaactctcg catccagtga gaactaccat acaccacatt acagctngga atgtntcca 120
 aatgtctggt caaatgatac aatggaacca ttcaatctta cacatgcacg aaagaacaag 180
 cgcttttgac atacaatgca caaaaaaaaa aggggggggg gaccacatgg attaaaattt 240
 taagtactca tcacatacat taagacacag ttctagtcca gtcnaaaatc agaactgcnt 300
 tgaaaaattt catgtatgca atccaaccaa agaacttnat tggatgatcat gantnctota 360
 ctacatcnac cttgatcatt gccaggaacn aaaagttnaa ancacncngt acaaaaaanaa 420
 tctgtaattn anttcaacct ccgtacngaa aaatnttntt tatacactcc c 471

<210> 191
 <211> 402
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(402)
 <223> n = A,T,C or G

<400> 191
 gagggattga aggtctgttc tastgtcggm ctgttcagcc accaactcta acaagttgct 60
 gtcttccact cactgtctgt aagcttttta acccagacwg tatcttcata aatagaacaa 120
 attcttcacc agtcacatct tctaggacct ttttgattc agttagtata agctcttcca 180
 ctccctttgt taagacttca tctggtaaag tcttaagttt tgtagaaagg aattyaattg 240
 ctggttctct aacaatgtcc tctccttgaa gtatttggct gaacaacca cctaaagtcc 300
 ctttgtgcat ccatttttaa tatacttaat agggcattgk tncactaggt taaattctgc 360
 aagagtcac tgtctgcaaa agttgcgtta gtatatctgc ca 402

<210> 192
 <211> 601
 <212> DNA

<400> 194

```

gaacggctgg accttgccct gcattgtgct tgctggcagg gaataccttg gcaagcagyt      60
ccagtccgag cagccccaga ccgctgccgc ccgaagctaa gcctgcctct ggccctcccc      120
tccgcctcaa tgcagaacca gtagtgggag cactgtgttt agagttaaga gtgaacactg      180
tttgatttta cttgggaatt tcctctgtta tatagctttt cccaatgcta atttccaaac      240
aacaacaaca aaataacatg tttgcctgtt aagttgtata aaagtaggtg attctgtatt      300
taaagaaaat attactgtta catatactgc ttgcaatttc tgtattttatt gktnctstgg      360
aaataaatat agttattaaa ggttgtcant cc                                     392

```

<210> 195

<211> 502

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(502)

<223> n = A,T,C or G

<400> 195

```

ccsttkgagg ggtkaggkyc cagttyccga gtggaagaaa caggccagga gaagtgcgtg      60
ccgagctgag gcagatgttc ccacagtgc cccagagacc stgggstata gtytctgacc      120
cctcncaagg aaagaccacs ttctggggac atgggctgga gggcaggacc tagaggcacc      180
aagggaaagg cccattccgg ggstgttccc cgaggaggaa ggggaaggggc tctgtgtgcc      240
ccccasgagg aagaggccct gagtccctgg atcagacacc ccttcacgtg tatccccaca      300
caaatgcaag ctcaccaagg tccccctca gtccccttcc stacaccctg amcggccact      360
gscscacacc caccagagc acgccacccg ccatggggar tgtgctcaag gartcgcnng      420
gcarcgtgga catctngtcc cagaaggggg cagaatctcc aatagangga ctgarcmsst      480
gctnanaaaa aaaaanaaaa aa                                     502

```

<210> 196

<211> 665

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(665)

<223> n = A,T,C or G

<400> 196

```

ggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc      60
cctctggaag ccttgccgag agcggacttt gtaattgttg gagaataact gctgaatttt      120
wagctgtttk gagttgatts gcaccactgc acccacaact tcaatatgaa aacyawttga      180
actwatthtat tatcttgtga aaagtataac aatgaaaatt ttgttcatac tgtattkac      240
aagtatgatg aaaagcaawa gatatatatt cttttattat gttaaaattat gattgccatt      300
attaatcggc aaaatgtgga gtgtatgttc ttttcacagt aatatatgcc ttttgtaact      360
tcacttgggt atttttattgt aaatgartta caaaattctt aatttaagar aatgggtatgt      420
watatthtat tcattaattt ctttcctkgt ttacgtwaat tttgaaaaga wtgcatgatt      480
tcttgacaga aatcgatctt gatgctgtgg aagtagtttg acccacatcc ctatgagttt      540
ttcttagaat gtataaagggt tgtagcccat cnaacttcaa agaaaaaaat gaccacatac      600
tttgcaatca ggctgaaatg tggcatgctn ttctaattcc aactttataa actagcaaan      660
aagtg                                              665

```

<210> 197

<211> 492
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(492)
 <223> n = A,T,C or G

<400> 197
 tttntttttt ttttttttgc aggaaggatt ccattttattg tggatgcatt ttcacaatat 60
 atgtttattg gagegatcca ttatcagtga aaagtatcaa gtgtttataa natttttagg 120
 aaggcagatt cacagaacat gctngtcngc ttgcagtttt acctcgtana gatnacagag 180
 aattatagtc naaccagtaa acnaggaatt tacttttcaa aagattaaat ccaaactgaa 240
 caaaattcta ccctgaaact tactccatcc aaatattgga ataanagtca gcagtgatac 300
 attctcttct gaactttaga ttttctagaa aaatatgtaa tagtgatcag gaagagctct 360
 tgttcaaaag tacaacnaag caatgttccc ttaccatagg ccttaattca aactttgatc 420
 catttcactc ccatcacggg agtcaatgct acctgggaca cttgtatttt gtcatnctg 480
 ancntggctt aa 492

<210> 198
 <211> 478
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(478)
 <223> n = A,T,C or G

<400> 198
 tttnttttgn atttcantct gtannaanta ttttcattat gtttattana aaaatatnaa 60
 tgtntccacn acaaatcatn ttacntnagt aagaggccan ctacattgta caacatacac 120
 tgagtatatt ttgaaaagga caagtttaaa gtanacncat attgccganc atancacatt 180
 tatacatggc ttgattgata tttagcacag canaaactga gtgagttacc agaaanaaat 240
 natatatgtc aatcngattt aagatacaaa acagatccta tggtagatan catcntgtag 300
 gagttgtggc tttatgttta ctgaaagtca atgcagttcc tgtacaaaaga gatggccgta 360
 agcattctag tacctctact ccattggttaa gaatcgtaca cttatgttta catatgtnc 420
 gggtaagaat tgtgttaagt naanttatgg agaggtccan gagaaaaatt tgatncaa 478

<210> 199
 <211> 482
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(482)
 <223> n = A,T,C or G

<400> 199
 agtgacttgt cctccaacaa aacccttga tcaagtttgt ggcactgaca atcagaccta 60
 tgctagttcc tgtcatctat tcgctactaa atgcagactg gaggggacca aaaaggggca 120
 tcaactocag ctggattatt ttggagcctg caaatctatt cctacttgta cggactttga 180

```

agtgattcag tttcctctac ggatgagaga ctgggtcaag aatatacctca tgcagcttta      240
tgaagccnac tctgaacacg ctgggttatct nagatgagaa ncagagaaat aaagtcnaga      300
aaatttacct ggangaaaag aggctttngg ctgggggacca tcccattgaa ccttctctta      360
anggacttta agaanaaaact accacatgtn tgtngtatcc tgggtgcengg ccgtttantg      420
aacntngacn ncaccttntt ggaatanant cttgaengcn tctgaactt gctcctctgc      480
ga                                                                                   482

```

```

<210> 200
<211> 270
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(270)
<223> n = A,T,C or G

```

```

<400> 200
cggccgcaag tgcaactcca gctggggccg tgcggacgaa gattctgcc a gcagttggtc      60
cgactgcgac gacggcgggc gcgacagtcg caggtgcagc gcgggcgcct ggggtcttgc      120
aaggctgagc tgacgccgca gaggtcgtgt cacgtcccac gaccttgacg ccgtcgggga      180
cagccggaac agagcccggg gaangcggga ggcctcgggg agccccctcg gaagggcggc      240
ccgagagata cgcaggtgca ggtggccgcc                                                                                   270

```

```

<210> 201
<211> 419
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(419)
<223> n = A,T,C or G

```

```

<400> 201
tttttttttt ttttggaaat tactgcgagc acagcaggtc agcaacaagt ttattttgca      60
gctagcaagg taacagggta gggcatggtt acatgttcag gtcaacttcc tttgtcgtgg      120
ttgattgggt tgtctttatg ggggcggggg ggggtagggg aaancgaagc anaantaaca      180
tggagtgggt gcaccctccc tgtagaacct gggtacnaaa gcttggggca gttcacctgg      240
tctgtgaccg tcattttctt gacatcaatg ttattagaag tcaggatatc ttttagagag      300
tccactgtnt ctggagggag attagggttt cttgccaana tccaancaa atccacntga      360
aaaagttgga tgatncangt acngaatacc ganggcatan ttctcatant cgggtggcca      419

```

```

<210> 202
<211> 509
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(509)
<223> n = A,T,C or G

```

```

<400> 202

```

009069"6225950

```

ttnttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt      60
tggcacttaa tccattttta tttcaaaatg tctacaaant ttnaatncnc cattatacng      120
gtnattttnc aaaatctaaa nnttattcaa atntnagcca aantccttac ncaaattnaa      180
tacnncaaa aatcaaaaat atacntntct ttcagcaaac ttngttacat aaattaaaaa      240
aatatatacg gctgggtgtt tcaaagtaca attatcttaa cactgcaaac atnttttnaa      300
ggaactaaaa taaaaaaaaa cactnccgca aagggttaaag ggaacaacaa attcntttta      360
caacancnnc nattataaaa atcatactct aaatccttag ggaatatata cttcacacng      420
ggatcttaac ttttactnca ctttgtttat ttttttanaa ccattgtntt gggcccaaca      480
caatggnaat nccnccnncn tggtactagt                                     509

```

```

<210> 203
<211> 583
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (583)
<223> n = A,T,C or G

```

```

<400> 203
tttttttttt ttttttttga cccccctctt ataaaaaaca agttaccatt ttatttttact      60
tacacatatt tattttataa ttggtattag atattcaaaa ggcagctttt aaaatcaaac      120
taaattggaaa ctgccttaga tacataattc ttaggaatta gcttaaaatc tgccataaagt      180
gaaaatcttc tctagctctt ttgactgtaa atttttgact cttgtaaaac atccaaattc      240
atttttcttg tctttaaaaa tatctaattc ttccattttt tccctattcc aagtcaattt      300
gcttctctag cctcatttcc tagctcttat ctactattag taagtggctt ttttctctaaa      360
agggaaaaca ggaagagana atggcacaca aaacaaacat tttatattca tatttctacc      420
tacgttaata aaatagcatt ttgtgaagcc agctcaaaag aaggcttaga tccttttatg      480
tccatttttag tactataaac atatcnaaag tgccagaatg caaaaagggtt gtgaacattt      540
attcaaaagc taatataaga tatttcacat actcatcttt ctg                                     583

```

```

<210> 204
<211> 589
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (589)
<223> n = A,T,C or G

```

```

<400> 204
ttttttttnt tttttttttt ttttttntct ttcttttttt ttganaatga ggatcgagtt      60
tttactctct tagatagggc atgaagaaaa ctcatctttc cagcttttaa ataacaatca      120
aatctcttat gctatatcat attttaagtt aaactaatga gtcactggct tatcttctcc      180
tgaaggaaat ctgttcattc ttctcattca tatagttata tcaagtacta ccttgcatat      240
tgagagggtt ttcttctcta ttacacata tatttccatg tgaatttgta tcaaaccctt      300
attttcatgc aaactagaaa ataatgtntt cttttgcata agagaagaga acaatatnag      360
cattacaaaa ctgctcaaat tgtttgttaa gnntatccat tataattagt tnggcaggag      420
ctaatacaaa tcacatttac ngacnagcaa taataaaact gaagtaccag ttaaataatcc      480
aaaataatta aggaacattt tttagcctgg gtataattag ctaattcact ttacaagcat      540
ttattnagaa tgaattcaca tggtattatt ccntagccca acacaatgg                                     589

```

```
<220>
<221> misc_feature
<222> (1)...(545)
<223> n = A,T,C or G
```

```
<210> 206
<211> 487
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(487)
<223> n = A,T,C or G
```

```
<210> 207
<211> 332
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(332)
<223> n = A,T,C or G
```

<400> 207

```

tgaattggct aaaagactgc atttttanaa ctagcaactc ttattttcttt cttttaaaaa 60
tacatagcat taaatcccaa atcctattta aagacctgac agcttgagaa ggtcactact 120
gcatttatag gaccttctgg tggttctgct gttacntttg aantctgaca atccttgana 180
atctttgcat gcagaggagg taaaaggat tggattttca cagaggaana acacagcgca 240
gaaatgaagg ggccaggctt actgagcttg tccactggag ggctcatggg tgggacatgg 300
aaaagaaggc agcctaggcc ctggggagcc ca 332

```

```

<210> 208
<211> 524
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(524)
<223> n = A,T,C or G

```

```

<400> 208
agggcggtgt gcggaggggc ttactgtttt gtctcagtaa caataaatac aaaaagactg 60
gttgtgttcc ggccccatcc aaccacgaag ttgattttct ttgtgtgcag agtgactgat 120
tttaaaggac atggagcttg tcacaatgtc acaatgtcac agtgtgaagg gcacactcac 180
tcccgcgtga ttcacattta gcaaccaaca atagctcatg agtccatact tgtaaatact 240
tttggcagaa tacttnttga aacttgcaga tgataactaa gatccaagat atttcccaaa 300
gtaaatagaa gtgggtcata atattaatta cctgttcaca tcagcttcca tttacaagtc 360
atgagcccag aactgacat caaactaagc ccacttagac tctcaccac cagtctgtcc 420
tgtcatcaga caggaggctg tcaccttgac caaattctca ccagtcaatc atctatccaa 480
aaaccattac ctgatccact tccggtaatg caccaccttg gtga 524

```

```

<210> 209
<211> 159
<212> DNA
<213> Homo sapien

```

```

<400> 209
gggtgaggaa atccagagtt gccatggaga aaattccagt gtcagcattc ttgctccttg 60
tggccctctc ctacactctg gccagagata ccacagtcaa acctggagcc aaaaaggaca 120
caaaggactc tcgacccaaa ctgccccaga ccctctcca 159

```

```

<210> 210
<211> 256
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(256)
<223> n = A,T,C or G

```

```

<400> 210
actccctggc agacaaaggc agaggagaga gctctgttag ttctgtgttg ttgaactgcc 60
actgaatttc tttccacttg gactattaca tgccanttga gggactaatg gaaaaacgta 120
tggggagatt ttanccaatt tangtntgta aatggggaga ctggggcagg cgggagagat 180
ttgcaggggt naaatgggan ggctggtttg ttanatgaac agggacatag gaggtaggca 240
ccaggatgct aaatca 256

```


<210> 211
 <211> 264
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(264)
 <223> n = A,T,C or G

<400> 211
 acattgtttt tttgagataa agcattgaga gagctctcct taacgtgaca caatggaagg 60
 actggaacac ataccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120
 atattcaagc acatatgtta tatattattc agttccatgt ttatagccta gttaaggaga 180
 ggggagatac attcngaaag aggactgaaa gaaatactca agtnggaaaa cagaaaaaga 240
 aaaaaaggag caaatgagaa gcct 264

<210> 212
 <211> 328
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(328)
 <223> n = A,T,C or G

<400> 212
 acccaaaaaat ccaatgctga atatttggct tcattattcc canattcttt gattgtcaaa 60
 ggatttaatg ttgtctcagc ttgggcactt cagttaggac ctaaggatgc cagccggcag 120
 gtttatatat gcagcaacaa tattcaagcg cgacaacagg ttattgaact tgcccggcag 180
 ttnaatttca ttcccattga cttgggatcc ttatcatcag ccagagagat tgaaaattta 240
 cccctacnac tctttactct ctgganaggg ccagtgggtg tagctataag cttggccaca 300
 ttttttttct ctttattcct ttgtcaga 328

<210> 213
 <211> 250
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(250)
 <223> n = A,T,C or G

<400> 213
 attatgagc agagcgacat atccnagtgt agactgaata aaactgaatt ctctccagtt 60
 taaagcattg ctactgaag ggatagaagt gactgccagg agggaaagta agccaaggct 120
 cattatgcca aagganatat acatttcaat tctccaaact tcttcctcat tccaagagtt 180
 ttcaatattt gcatgaacct gctgataanc catgttaana aacaaatatt tctctnaact 240
 tctcatoggt 250

<210> 214

<210> 217
 <211> 262
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(262)
 <223> n = A,T,C or G

<400> 217
 acctacgtgg gtaagtttan aaatgttata atttcaggaa naggaacgca tataattgta 60
 tcttgcctat aattttctat ttttaataagg aaatagcaaa ttgggggtggg gggaatgtag 120
 ggcatcttac agtttgagca aaatgcaatt aaatgtggaa ggacagcact gaaaaatttt 180
 atgaataatc tgtatgatta tatgtctcta gagtagattt ataattagcc acttacccta 240
 atatccttca tgcttgtaaa gt 262

<210> 218
 <211> 205
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(205)
 <223> n = A,T,C or G

<400> 218
 accaaggtgg tgcattaccg gaantggatc aangacacca tcgtggccaa cccctgagca 60
 cccctatcaa ctcccttttg tagtaaaactt ggaaccttgg aaatgaccag gccaaagactc 120
 aggctctccc agttctactg acctttgtcc ttangtntna ngtccagggt tgctaggaaa 180
 anaaatcagc agacacaggt gtaaa 205

<210> 219
 <211> 114
 <212> DNA
 <213> Homo sapien

<400> 219
 tactgttttg tctcagtaac aataaaatata aaaagactgg ttgtgttccg gccccatcca 60
 accacgaagt tgatttctct tgtgtgcaga gtgactgatt ttaaaggaca tgga 114

<210> 220
 <211> 93
 <212> DNA
 <213> Homo sapien

<400> 220
 actagccagc acaaaaggca gggtagcctg aattgctttc tgctctttac atttctttta 60
 aaataagcat ttagtgctca gtccttactg agt 93

<210> 221
 <211> 167

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(167)
<223> n = A,T,C or G

<400> 221
actangtgca ggtg'gcaca aatatttgtc gatattccct tcatcttgga ttccatgagg 60
tcttttgccc agcctgtggc tctactgtag taagtttctg ctgatgagga gccagnatgc 120
ccccactac cttccctgac gctccccana aatcacccaa cctctgt 167

<210> 222
<211> 351
<212> DNA
<213> Homo sapien

<400> 222
agggcggtgt ggcgagggcg gtactgacct cattagtagg aggatgcatt ctggcacccc 60
gttcttcacc tgtcccccaa tccttaaaaag gccatactgc ataaagtcaa caacagataa 120
atgtttgtctg aattaaagga tggatgaaaa aaattaataa tgaatttttg cataatccaa 180
ttttctcttt tatatttcta gaagaagttt ctttgagcct attagatccc gggaatcttt 240
taggtgagca tgattagaga gctttaggt tgcttttaca tatactctggc atatttgagt 300
ctcgtatcaa aacaatagat tggtaaaggt ggtattattg tattgataag t 351

<210> 223
<211> 383
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(383)
<223> n = A,T,C or G

<400> 223
aaaacaaaca aacaaaaaaa acaattcttc attcagaaaa attatcttag ggactgatat 60
tggttaattat ggtcaattta atwrtrttkt ggggcatttc cttacattgt cttgacaaga 120
ttaaagtgtc tgtgccaaaa ttttgtattt tatttgagga cttcttatca aaagtaatgc 180
tgccaaagga agtctaagga attagtagtg ttcccmtcac ttgtttggag tgtgctattc 240
taaaagattt tgatttctctg gaatgacaat tatattttta ctttggtggg ggaaanagtt 300
ataggaccac agtcttcact tctgatactt gtaaattaat cttttattgc acttggtttg 360
accattaagc tatatgttta aaa 383

<210> 224
<211> 320
<212> DNA
<213> Homo sapien

<400> 224
cccctgaagg cttcttggtta gaaaatagta cagttacaac caataggaac aacaaaaaga 60
aaaagtttgt gacattgtag tagggagtgt gtacccttta ctcccatca aaaaaaaat 120
ggatacatgg ttaaaggata raagggcaat attttatcat atgttctaaa agagaaggaa 180

gagaaaatac tacttttctcr aaatggaagc ccttaaaggt gctttgatac tgaaggacac 240
 aaatgtggcc gtccatcctc ctttaragtt gcatgacttg gacacggtaa ctgttgacgt 300
 tttaractcm gcatgttgac 320

<210> 225
 <211> 1214
 <212> DNA
 <213> Homo sapien

<400> 225
 gaggactgca gcccgactc gcagccctgg caggcggcac tggatcatgga aaacgaattg 60
 ttctgctcgg gcgtcctggt gcatccgcag tgggtgctgt cagccgcaca ctgtttccag 120
 aactcctaca ccatcgggct gggcctgcac agtcttgagg ccgaccaaga gccagggagc 180
 cagatgggtg aggccagcct ctccgtacgg caccagagt acaacagacc cttgctcgtc 240
 aacgacctca tgctcatcaa gttggacgaa tccgtgtccg agtctgacac catccggagc 300
 atcagcattg cttcgcagtg ccctaccgcg ggggaactctt gcctcgtttc tggctgggggt 360
 ctgctggcga acggcagaat gcctaccgtg ctgcagtgcg tgaacgtgtc ggtgggtgtc 420
 gaggaggtct gcagtaagct ctatgacccg ctgtaccacc ccagcatgtt ctgcgccggc 480
 ggagggcaag accagaagga ctctgcaac ggtgactctg gggggccctt gatctgcaac 540
 ggggtacttg agggccttgt gtctttcggg aaagccctgt gtggccaagt tggcgtgcca 600
 ggtgtctaca ccaacctctg caaattcact gagggtgatag agaaaaccgt ccaggccagt 660
 taactctggg gactgggaac ccatgaaatt gacccccaaa tacatcctgc ggaaggaatt 720
 caggaatata tgttcccagc cctcctccc tcaggcccag gaggccaggc cccagccccc 780
 tctcctctca aaccaagggt acagatcccc agccctcct cctcagacc caggagtcca 840
 gacccccag cctcctccc ctcagaccca ggagtccagc cctcctccc tcagaccag 900
 gaggccagac ccccgagccc ctctcctc agacccaggg gtccaggccc ccaacccctc 960
 ctccctcaga ctcagaggtc caagccccc acccctcctt cccagacccc agagggtccag 1020
 gtccagccc ctctcctc agacccagcg gtccaatgcc acctagactc tccctgtaca 1080
 cagtcccc ttgtggcag ttgacccaac cttaccagtt ggtttttcat tttttgtccc 1140
 tttccctag atccagaaat aaagtctaag agaagcgcaa aaaaaaaaaa aaaaaaaaaa 1200
 aaaaaaaaaa aaaa 1214

<210> 226
 <211> 119
 <212> DNA
 <213> Homo sapien

<400> 226
 accagtatg tgcagggaga cggaacccca tgtgacagcc cactccacca gggttcccaa 60
 agaacctggc ccagtcataa tcattcatcc tgacagtggc aataatcacg ataaccagt 119

<210> 227
 <211> 818
 <212> DNA
 <213> Homo sapien

<400> 227
 acaattcata gggacgacca atgaggacag ggaatgaacc cggctctccc ccagccctga 60
 tttttgtac atatggggtc ctttttcatt ctttgcaaaa aactgggtt ttctgagAAC 120
 acggacgggt cttagcaca tttgtgaaat ctgtgtaraa ccgggctttg caggggagat 180
 aattttctc ctctggagga aaggtggtga ttgacaggca gggagacagt gacaaggcta 240
 gagaaagcca cgctcgccct tctctgaacc aggatggaac ggcagacccc tgaaaacgaa 300
 gcttgctccc ttccaatcag ccacttctga gaaccccat ctaacttct actggaaaag 360
 agggcctcct caggagcagt ccaagagttt tcaaagataa cgtgacaact accatctaga 420

ggaaaggggtg caccctcagc agagaagccg agagcttaac tctggtcggt tccagagaca 480
 acctgctggc tgtcttgga tgcgcccagc ctttgagagg ccactacccc atgaacttct 540
 gccatccact ggacatgaag ctgaggacac tgggcttcaa cactgagttg tcatgagagg 600
 gacaggctct gccctcaagc cggtcgaggg cagcaaccac tctcctcccc tttctcacgc 660
 aaagccattc ccacaaatcc agaccatacc atgaagcaac gagacccaaa cagtttgggt 720
 caagaggata tgaggactgt ctcagcctgg ctttgggctg acaccatgca cacacacaag 780
 gtccacttct aggttttcag cctagatggg agtcgtgt 818

<210> 228

<211> 744

<212> DNA

<213> Homo sapien

<400> 228

actggagaca ctgttgaact tgatcaagac ccagaccacc ccaggctctc ttcgtgggat 60
 gtcattgacgt ttgacatacc tttggaacga gcctcctcct tggaagatgg aagaccgtgt 120
 tctgtggccga cctggcctct cctggcctgt ttcttaagat gcggagtcac atttcaatgg 180
 taggaaaagt ggcttcgtaa aatagaagag cagtcactgt ggaactacca aatggcgaga 240
 tgctcgggtg acattggggg gctttgggat aaaagattta tgagccaact attctctggc 300
 accagattct aggccagttt gttccactga agcttttccc acagcagtc acctctgcag 360
 gctggcagct gaatggcttg ccggtggctc tgtggcaaga tcacactgag atcgatgggt 420
 gagaaggcta ggatgcttgt ctagtgttct tagctgtcac gttggctcct tccaggttgg 480
 ccagacgggtg ttggccactc cttctaaaaa cacaggcgcc ctctcgttga cagtgacccg 540
 ccgtgggtatg ctttgcccca ttccagcagt cccagttatg catttcaagt ttggggtttg 600
 ttcttttctg taatgttctt ctgtgttgtc agctgtcttc atttctcggg ctaagcagca 660
 ttgggagatg tggaccagag atccactcct taagaaccag tggcgaaaga cactttcttt 720
 cttcactctg aagtagctgg tggg 744

<210> 229

<211> 300

<212> DNA

<213> Homo sapien

<400> 229

cgagtctggg ttttgtctat aaagtttgat ccctcctttt ctcatccaaa tcatgtgaac 60
 cattacacat cgaaataaaa gaaaggtggc agacttgccc aacgccaggc tgacatgtgc 120
 tgcagggttg ttgtttttta attattattg ttagaaacgt caccacagc ccctgttaat 180
 ttgtatgtga cagccaactc tgagaaggtc ctatttttcc acctgcagag gatccagtct 240
 cactaggctc ctcttgccc tcacactgga gtctccgcca gtgtgggtgc ccactgacat 300

<210> 230

<211> 301

<212> DNA

<213> Homo sapien

<400> 230

cagcagaaca aatacaaata tgaagagtgc aaagatctca taaaatctat gctgaggaat 60
 gagcgacagt tcaaggagga gaagcttgca gagcagctca agcaagctga ggagctcagg 120
 caatataaag tcttggttca cactcaggaa cgagagctga cccagttaag ggagaagttg 180
 cggaaggga gagatgcctc cctctcattg aatgagcatc tccaggccct cctcactccg 240
 gatgaaccgg acaagtccca ggggcaggac ctccaagaaa cagacctcgg ccgcgaccac 300
 g 301

<210> 231

<211> 301
 <212> DNA
 <213> Homo sapien

<400> 231
 gcaagcacgc tggcaaactct ctgtcaggtc agctccagag aagccattag tcatttttagc 60
 caggaactcc aagtccacat ccttggaac tggggacttg cgcaggttag ccttgaggat 120
 ggcaacacgg gactttctcat caggaagtgg gatgtagatg agctgatcaa gacggccagg 180
 tctgaggatg gcaggatcaa tgatgtcagg ccggttggtta ccgccaatga tgaacacatt 240
 tttttttgtg gacatgccat ccattttctgt caggatctgg ttgatgactc ggtcagcagc 300
 c 301

<210> 232
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 232
 agtaggtatt tcgtgagaag ttcaacacca aaactggaac atagttctcc ttcaagtgtt 60
 ggcgacagcg gggcttcctg attctggaat ataactttgt gtaaattaac agccacctat 120
 agaagagtcc atctgctgtg aaggagagac agagaactct gggttccgtc gtccctgtcca 180
 cgtgctgtac caagtgtctg tgccagcctg ttacctgttc tcaactgaaa tctggctaata 240
 gctcttgtgt atcacttctg attctgacaa tcaatcaatc aatggcctag agcactgact 300
 g 301

<210> 233
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 233
 atgactgact tcccagtaag gctctctaag gggtaagtag gaggatccac aggatttgag 60
 atgctaaggc cccagagatc gtttgatcca accctcttat tttcagaggg gaaaatgggg 120
 cctagaagtt acagagcatc tagctgggtg gctggcacc ctggcctcac acagactccc 180
 gagtagctgg gactacaggc acacagtcac tgaagcaggc cctgttagca attctatgag 240
 taaaaattaa catgagatga gtagagactt tattgagaaa gcaagagaaa atcctatcaa 300
 c 301

<210> 234
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 234
 aggtcctaca catcgagact catccatgat tgatatgaat ttaaaaaatta caagcaaaga 60
 catttttattc atcatgatgc tttcttttgt ttcttctttt cgttttcttc tttttctttt 120
 tcaatttcag caacatactt ctcaatttct tcaggattta aaatcttgag ggattgatct 180
 cgctcatga cagcaagttc aatgtttttg ccacctgact gaaccacttc caggagtgcc 240
 ttgatcacca gcttaatggg cagatcatct gcttcaatgg cttcgtcagt atagttcttc 300
 t 301

<210> 235
 <211> 283
 <212> DNA

<213> Homo sapien

<400> 235

tggggctgtg	catcaggcgg	gtttgagaaa	tattcaattc	tcagcagaag	ccagaatttg	60
aattccctca	tcttttaggg	aatcatttac	caggtttgga	gaggattcag	acagctcagg	120
tgctttcact	aatgtctctg	aacttctgtc	cctctttgtt	catggatagt	ccaataaata	180
atgttatctt	tgaactgatg	ctcataggag	agaatataag	aactctgagt	gatatcaaca	240
ttagggattc	aaagaaatat	tagatttaag	ctcacactgg	tca		283

<210> 236

<211> 301

<212> DNA

<213> Homo sapien

<400> 236

aggtcctcca	ccaactgcct	gaagcacggt	taaaattggg	aagaagtata	gtgcagcata	60
aatactttta	aatcgatcag	atttccctaa	cccacatgca	atcttcttca	ccagaagagg	120
tcggagcagc	atcattaata	ccaagcagaa	tcggtaatag	ataaatacaa	tggtatatag	180
tgggtagacg	gcttcagtag	tacagtgtac	tgtggtatcg	taatctggac	ttgggttgta	240
aagcatcgtg	taccagtcag	aaagcatcaa	tactcgacat	gaacgaatat	aaagaacacc	300
a						301

<210> 237

<211> 301

<212> DNA

<213> Homo sapien

<400> 237

cagtggtagt	ggtggtaggac	gtggcggttg	tcgtgggtgcc	ttttttggtg	cccgtcacaa	60
actcaatttt	tgttcgctcc	tttttgccct	tttccaattt	gtccatctca	atttttctggg	120
ccttggtctaa	tgccatcatag	taggagtcct	cagaccagcc	atgggggatca	aacatatacct	180
ttgggtagtt	ggtgccaagc	tcgtcaatgg	cacagaatgg	atcagcttct	cgtaaatacta	240
gggttcogaa	attcttttctt	cctttggata	atgtagttca	tatccattcc	ctccttttate	300
t						301

<210> 238

<211> 301

<212> DNA

<213> Homo sapien

<400> 238

gggcagggttt	ttttttttttt	ttttttgatg	gtgcagaccc	ttgcttttatt	tgtctgactt	60
gttcacagtt	cagccccctg	ctcagaaaac	caacggggcca	gctaaggaga	ggaggaggca	120
ccttgagact	tccggagtcg	aggctctcca	gggttcccca	gcccataaat	cattttctgc	180
acccccctgcc	tggaagcag	ctccctgggg	ggtgggaatg	ggtgactaga	agggatttca	240
gtgtgggacc	cagggtctgt	tcttcacagt	aggagggtga	agggatgact	aattttcttta	300
t						301

<210> 239

<211> 239

<212> DNA

<213> Homo sapien

<400> 239


```
<210> 240
<211> 300
<212> DNA
<213> Homo sapien
```

```
<210> 241
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 242
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 243
<211> 301
<212> DNA
<213> Homo sapien
```

<400>	243						
aggtaagtcc	cagtttgaag	ctcaaaagat	ctggtatgag	cataggctca	tgcacgacat		60
ggtggcccaa	gctatgaaat	cagagggagg	cttcactctg	gctgtaaaa	actatgatgg		120
tgacgtgcag	tgggactctg	tggcccaagg	gtatggctct	ctcggcatga	tgaccagcgt		180
gctggtttgt	ccagatggca	agacagtaga	agcagaggct	gccacggga	ctgtaacccg		240
tcactaccgc	atgttccaqa	aaggacagga	gacgtccacc	aatccattg	cttccatttt		300

<213> Homo sapien

<400> 252

```
gcaaccaatc actctgtttc acgtgacttt tatcaccata caatttgtgg catttcctca    60
ttttctacat tgtagaatca agagtgtaaa taaatgtata tcgatgtctt caagaatata    120
tcatttcctt ttacttagga acccattcaa aatataagtc aagaatctta atatcaacaa    180
atatatcaag caaactggaa ggcagaataa ctaccataat ttagtataag taccctaaagt    240
tttataaatc aaaagcccta atgataacca tttttagaat tcaatcatca ctgtagaatc    300
a                                     301
```

<210> 253

<211> 301

<212> DNA

<213> Homo sapien

<400> 253

```
ttccctaaga agatgttatt ttgttgggtt ttgttcccc tccatctcga ttctcgtacc    60
caactaaaaa aaaaaataa agaaaaaatg tgctgcgttc tgaaaaataa ctccttagct    120
tggtctgatt gttttcagac cttaaaatat aaacttgttt cacaagcttt aatccatgtg    180
gatttttttt cttagagaac cacaaaacat aaaaggagca agtcggactg aatacctgtt    240
tccatagtgc ccacagggtta ttcttcacat tttctccata ggaaaatgct ttttcccaag    300
g                                     301
```

<210> 254

<211> 301

<212> DNA

<213> Homo sapien

<400> 254

```
cgctgcgcct ttcccttggg ggaggggcaa ggccagaggg ggtccaagtg cagcacgagg    60
aacttgacca attcccttga agcgggtggg ttaaaccctg taaatgggaa caaatcccc    120
ccaaatctct tcatcttacc ctggtggact cctgactgta gaattttttg gttgaaacaa    180
gaaaaaaata agcctttgga cttttcaagg ttgcttaaca ggtactgaaa gactggcctc    240
acttaaaactg agccaggaaa agctgcagat ttattaatgg gtgtgttagt gtgcagtgcc    300
t                                     301
```

<210> 255

<211> 302

<212> DNA

<213> Homo sapien

<400> 255

```
agcttttttt tttttttttt tttttttttt ttcattaaaa aatagtgtctc tttattataa    60
attactgaaa tgtttctttt ctgaatataa atataaatat gtgcaaagtt tgacttggat    120
tggtattttt ttgagttctt caagcatctc ctaataccct caagggcctg agtagggggg    180
aggaaaaagg actggaggtg gaatctttat aaaaaacaag agtgattgag gcagattgta    240
aacattatta aaaaacaaga aacaaacaaa aaaatagaga aaaaaaccac cccaacacac    300
aa                                     302
```

<210> 256

<211> 301

<212> DNA

<213> Homo sapien

000000"6225950

<220>
 <221> misc_feature
 <222> (1)...(301)
 <223> n = A,T,C or G

<400> 256
 gttccagaaa acattgaagg tggcttccca aagtctaact agggataccc cctctagcct 60
 aggaccctcc tccccacacc tcaatccacc aaaccatcca taatgcaccc agataggccc 120
 acccccacaaa gcctggacac cttgagcaca cagttatgac caggacagac tcattcttat 180
 aggcaaatag ctgctggcaa actggcatta cctgggttgt ggggatgggg gggcaagtgt 240
 gtggcctctc ggctgggta gcaagaacat tcagggtagg cctaagttan tcgtgttagt 300
 t 301

<210> 257
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 257
 gttgtggagg aactctggct tgctcattaa gtcctactga ttttactat cccctgaatt 60
 tccccactta tttttgtctt tcactatcgc aggccttaga agaggtctac ctgcctccag 120
 tcttacctag tccagtctac cccctggagt tagaatggcc atcctgaagt gaaaagtaat 180
 gtcacattac tcccttcagt gatttcttgt agaagtgcc atccctgaat gccaccaaga 240
 tcttaattct cactcttcta atcttatctc tttgactcct ctttacaccg gagaaggctc 300
 c 301

<210> 258
 <211> 301
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(301)
 <223> n = A,T,C or G

<400> 258
 cagcagtagt agatgccgta tgccagcacg cccagcactc ccaggatcag caccagcacc 60
 aggggcccag ccaccaggcg cagaagcaag ataaacagta ggctcaagac cagagccacc 120
 cccagggcaa caagaatcca ataccaggac tgggcaaaat cttcaaagat cttaacactg 180
 atgtctcggg cattgaggct gtcaataana cgctgatccc ctgctgtatg gtggtgtcat 240
 tgggtgatccc tgggagcgcc ggtggagtaa cgttgggtcca tggaaagcag cgcccacaac 300
 t 301

<210> 259
 <211> 301
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(301)
 <223> n = A,T,C or G

009060" 62245960

<400> 259

```

tcatatatgc aaacaaatgc agactangcc tcaggcagag actaaaggac atctcttggg      60
gtgtcctgaa gtgatttgga cccctgaggg cagacaccta agtaggaatc ccagtgggaa      120
gcaaagccat aaggaagccc aggattcctt gtgatcagga agtgggcccag gaaggctctgt      180
tccagctcac atctcatctg catgcagcac ggaccggatg cggccactgg gtcttggctt      240
ccctcccatc ttctcaagca gtgtccttgt tgagccattt gcateccttg ctccagggtg      300
c                                                                    301

```

<210> 260

<211> 301

<212> DNA

<213> Homo sapien

<400> 260

```

ttttttttct ccctaaggaa aaagaaggaa caagtctcat aaaaccaaat aagcaatggt      60
aagggtgtctt aacttgaaaa agattaggag tcaactgggtt acaagttata attgaatgaa      120
agaactgtaa cagccacagt tggccatttc atgccaatgg cagcaaacia caggattaac      180
tagggcaaaa taaataagtg tgtggaagcc ctgataagtg cttataaac agactgatcc      240
actgagacat cagtacctgc ccgggcggcc gctcgagccg aattctgcag atatccatca      300
c                                                                    301

```

<210> 261

<211> 301

<212> DNA

<213> Homo sapien

<400> 261

```

aaatattcga gcaaatcctg taactaatgt gtctccataa aaggctttga actcagtga      60
tctgcttcca tccacgattc tagcaatgac ctctcggaca tcaaagctcc tcttaagggtt      120
agcaccaact attccataca attcatcagc aggaaataaa ggctcttcag aaggttcaat      180
ggtgacatcc aatttcttct gataatttag attcctcaca accttcctag ttaagtgaag      240
ggcatgatga tcatccaaag cccagtggtc acttactcca gactttctgc aatgaagatc      300
a                                                                    301

```

<210> 262

<211> 301

<212> DNA

<213> Homo sapien

<400> 262

```

gaggagagcc tgttacagca tttgtaagca cagaatactc caggagtatt tgtaattgtc      60
tgtgagcttc ttgccgcaag tctctcagaa atttaaaaag atgcaaatcc ctgagtcacc      120
cctagacttc ctaaaccaga tcctctgggg ctggaacctg gcactctgca tttgtaatga      180
gggctttctg gtgcacacct aattttgtgc atctttgccc taaatcctgg attagtcccc      240
catcattacc cccacattat aatgggatag attcagagca gatactctcc agcaagaat      300
c                                                                    301

```

<210> 263

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 263

```

tttagcttgt ggtaaatgac tcacaaaact gatttttaaaa tcaagttaat gtgaattttg    60
aaaattacta cttaatccta attcacaata acaatggcat taaggtttga cttgagttgg    120
ttcttagtat tatttatggt aaataggctc ttaccacttg caaataactg gccacatcat    180
taatgactga cttcccagta aggctctcta aggggtaagt angaggatcc acaggatttg    240
agatgctaag gcccagaga tcgtttgatc caaccctctt attttcagag gggaaaatgg    300
g                                                                    301

```

<210> 264

<211> 301

<212> DNA

<213> Homo sapien

<400> 264

```

aaagacgtta aaccactcta ctaccacttg tggaactctc aaagggtaaa tgacaaaascc    60
aatgaatgac tctaaaaaca atatttacat ttaatggttt gtagacaata aaaaaacaag    120
gtggatagat ctagaattgt aacatttttaa gaaaaccata scatttgaca gatgagaaaag    180
ctcaattata gatgcaaagt tataactaaa ctactatagt agtaaagaaa tacatttcac    240
acccttcata taaattcact atcttggctt gaggcactcc ataaaatgta tcacgtgcat    300
a                                                                    301

```

<210> 265

<211> 301

<212> DNA

<213> Homo sapien

<400> 265

```

tgcccaagtt atgtgtaagt gtatccgcac ccagaggtaa aactacactg tcattcttgt    60
cttcttgtga cgcagtattt cttctctggg gagaagccgg gaagtcttct cctggctcta    120
catattcttg gaagtctcta atcaactttt gttccatttg tttcatttct tcaggaggga    180
ttttcagttt gtcaacatgt tctctaacaa cacttgccca tttctgtaaa gaatccaaag    240
cagtccaagg ctttgacatg tcaacaacca gcataactag agtatccttc agagatacgg    300
c                                                                    301

```

<210> 266

<211> 301

<212> DNA

<213> Homo sapien

<400> 266

```

taccgtctgc ctttctctcc atccaggcca tctgcgaatc tacatgggtc ctctatttcg    60
acaccagatc actctttcct ctaccacacag gcttgctatg agcaagagac acaacctcct    120
ctcttctgtg ttccagcttc ttttctgttt cttccacccc cttaagttct attcctgggg    180
atagagacac caatacccat aacctctctc ctaagcctcc ttataaccca ggggtgcacag    240
cacagactcc tgacaactgg taaggccaat gaactgggag ctcacagctg gctgtgcctg    300
a                                                                    301

```

<210> 267

<211> 301

<212> DNA

<213> Homo sapien

<400> 267

```

aaagagcaca ggccagctca gcctgccctg gccatctaga ctcagcctgg ctccatgggg      60
gttctcagtg ctgagtccat ccaggaaaag ctcacctaga ctttctgagg ctgaatcttc      120
atcctcacag gcagcttctg agagcctgat attcctagcc ttgatgggtct ggagtaaagc      180
ctcattctga ttctctcct tcttttcttt caagttggct ttcttcacat ccctctgttc      240
aattcgcttc agcttgctctg ctttagccct catttcacga agcttcttct ctttggcatc      300
t                                                                 301

```

<210> 268

<211> 301

<212> DNA

<213> Homo sapien

<400> 268

```

aatgtctcac tcaactactt cccagcctac cgtggcctaa ttctgggagt tttcttctta      60
gatcttggga gagctgggtc ttctaaggag aaggaggaag gacagatgta actttggatc      120
tcgaagagga agtctaattg aagtaattag tcaacgggtcc ttgttttagac tcttgggaata      180
tgctgggtgg ctcagtgagc ccttttggag aaagcaagta ttattcttaa ggagtaacca      240
cttcccattg ttctactttc taccatcatc aattgtatat tatgtattct ttggagaact      300
a                                                                 301

```

<210> 269

<211> 301

<212> DNA

<213> Homo sapien

<400> 269

```

taacaatata cactagctat ctttttaact gtccatcatt agcaccaatg aagattcaat      60
aaaattacct ttattcacac atctcaaaac aattctgcaa attcttagtg aagtttaact      120
atagtcacag accttaaata ttcacattgt tttctatgtc tactgaaaat aagttcacta      180
cttttctgga tattctttac aaaatcttat taaaattcct ggtattatca cccccaatta      240
tacagtagca caaccacctt atgtagtttt tacatgatag ctctgtagaa gtttcacatc      300
t                                                                 301

```

<210> 270

<211> 301

<212> DNA

<213> Homo sapien

<400> 270

```

cattgaagag cttttgcgaa acatcagaac acaagtgcct ataaaaattaa ttaagcctta      60
cacaagaata catattcctt ttatttctaa ggagttaaac atagatgtag ctgatgtgga      120
gagcttgctg gtgcagtgca tattggataa cactattcat ggccgaattg atcaagtcaa      180
ccaactcctt gaactggatc atcagaagaa ggggtggtgca cgatatactg cactagataa      240
tggaaccaacc aactaaattc tctcaccagg ctgtatcagt aaactggctt aacagaaaaac      300
a                                                                 301

```

<210> 271

<211> 301

<212> DNA

<213> Homo sapien

<220>

009959" 6225950

<400> 271

<210> 272

<211> 301

<212> DNA

<213> Homo sapien

<400> 272

taaattgcta	agccacagat	aacaccaatc	aaatggaaca	aatcactgtc	ttcaaatgtc	60
ttatcagaaa	accaaatgag	cctggaatct	tcataatacc	taaacatgcc	gtatttagga	120
tccaataaatt	ccctcatgat	gagcaagaaa	aattctttgc	gcacctctcc	tgcattccaca	180
gcatcttctc	caacaaatat	aaccttgagt	ggcttcttgt	aatctatgtt	ctttgttttc	240
ctaaggactt	ccattgcata	tctacaata	ttttctctac	gcaccactag	aattaagcag	300
g						301

<210> 273

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (301)$

<223> n = A, T, C or G

<400> 273

acatgtgtgt	atgtgtatct	ttgggaaaaan	aanaagacat	cttgttttayt	atTTTTTTTgg	60
agagangctg	ggacatggat	aatcacwtaa	tttgctayta	tyactTTta	ctgactytaa	120
gaaccgtcta	aaaataaaaat	ttaccatgtc	dtatatTCct	tatagtatgc	ttatttcacc	180
tyttttctgt	ccagagagag	tatcagtgac	ananatttma	gggtgaamac	atgmattggg	240
gggacttnty	tttacngagm	accctgcccc	sgcgccctcg	makcngant	ccgcsananc	300
t						301

<210> 274

<211> 301

<212> DNA

<213> Homo sapien

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (301)$

<223> n = A, T, C or G

<400> 274

cttatataact cttttctcaga ggcaaaagag gagatgggta atgtagacaa ttctttgagg 60
aacagtaaatt gattattaga gagaangaat ggaccaagga gacagaaatt aacttgtaaa 120
tgattctctt tgggaatctga atgagatcaa gagggcagct ttagcttggtg gaaaagtcca 180
tctaggtatg gttgcattct cgtcttcttt totgcagtag ataatgaggt aaccgaaggc 240
aattgtgctt cttttgataa gaagctttct tggtcatatc aggaaattcc aganaaaagtc 300
c 301

<210> 275
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (301)
<223> n = A,T,C or G

<400> 275
tcggtgtcag cagcacgtgg cattgaacat tgcaatgtgg agcccaaacc acagaaaatg 60
gggtgaaatt ggccaacttt ctattaactt atgttggaac ttttgccacc aacagtaagc 120
tggcccttct aataaaagaa aattgaaagg tttctcacta aacggaatta agtagtggag 180
tcaagagact cccaggcctc agcgtacctg cccgggcggc cgctcgaagc cgaattctgc 240
agatatccat cacactggcg gncgctcgan catgcatcta gaaggnccaa ttgcgcctat 300
a 301

<210> 276
<211> 301
<212> DNA
<213> Homo sapien

<400> 276
tgtacacata ctcaataaat aaatgactgc atttgtgtat tattactata ctgattatat 60
ttatcatgtg acttctaatt agaaaatgta tccaaaagca aaacagcaga tatacaaaat 120
taaagagaca gaagatagac attaacagat aaggcaactt atacattgag aatccaaatc 180
caatacattt aaacatttgg gaaatgaggg ggacaaatgg aagccagatc aaatttgtgt 240
aaaactattc agtatgtttc ccttgcttca tgtctgagaa ggctctcctt caatggggat 300
g 301

<210> 277
<211> 301
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (301)
<223> n = A,T,C or G

<400> 277
tttgttgatg tcagtatttt attacttgcg ttatgagtgc tcacctggga aattctaaag 60
atacagagga cttggaggaa gcagagcaac tgaatttaatt ttaaaagaag gaaaacattg 120
gaatcatggc actcctgata ctttcccaa tcaacactct caatgccccca cctcgtcct 180
caccatagtg gggagactaa agtggccacg gatttgctt angtgtgcag tgcgttctga 240
gttcnctgtc gattacatct gaccagtctc ctttttccga agtcctccg ttcaatcttg 300

301

taccactaca	ctccagcctg	ggcaacagag	caagacctgt	ctcaaagcat	aaaatggaat	60
aacatatcaa	atgaaacagg	gaaaatgaag	ctgacaatth	atggaagcca	gggcttgtca	120
cagtcctctac	tggtattatg	cattacctgg	gaatttatat	aagcccttaa	taataatgcc	180
aatgaacatc	tcattgtgtgc	tcacaatgtt	ctggcactat	tataagtgtc	tcacaggttt	240
tatgtgtttct	tcgtaactth	atggantagg	tactcggccg	cgaacacgct	aagccgaatt	300
c						301

<400> 279						
aaagcaggaa	tgacaaagct	tgctttttctg	gtatgtttcta	ggtgtatttgt	gactttttact	60
gttatattaa	ttgccaatat	aagtaaatat	agatttatata	tgtatagtgt	ttcacaaagc	120
ttagaccttt	accttcacag	cacccacacag	tgcttgatat	ttcagagtc	gtcattgggt	180
atacatgtgt	agttccaaag	cacataagct	agaanaanaa	atattttctag	ggagcactac	240
catctgtttt	cacatgaat	gccacacaca	tagaactcca	acatcaattt	cattgcacag	300
a						301

<400> 280							
ggtactggag	ttttcctccc	ctgtgaaaac	gtaactactg	ttggggagtga	attgagggatg		60
tagaaaggtg	gtggaaccaa	attgtgggtca	atggaaatag	gagaatatgg	ttctcactct		120
tgagaaaaaa	acctaagatt	agcccaggta	gttgcctgta	acttcagttt	ttctgcctgg		180
gtttgatata	gttttagggtt	ggggtttagat	taagatctaa	attacatcag	gacaaagaga		240
cagactatta	actccacagt	taattaagga	ggtatgttcc	atgtttattt	gttaaagcag		300
t							301

<210>	281
<211>	301
<212>	DNA

<213> Homo sapien

<400> 281

```
aggtacaaga aggggaatgg gaaagagctg ctgctgtggc attgttcaac ttgatattc    60
gccgagcaat ccaaatcctg aatgaagggg catcttctga aaaaggagat ctgaatctca    120
atgtggtagc aatggcttta tcgggttata cggatgagaa gaactccctt tggagagaaa    180
tgtgtagcac actgcgatta cagctaaata acccgatatt gtgtgtcatg tttgcatttc    240
tgacaagtga aacaggatct tacgatggag ttttgtatga aaacaaagtt gcagtacctc    300
g                                          301
```

<210> 282

<211> 301

<212> DNA

<213> Homo sapien

<400> 282

```
caggtactac agaattaaaa tactgacaag caagtagttt cttggcgtgc acgaattgca    60
tccagaaccc aaaaattaag aaattcaaaa agacattttg tgggcacctg ctagcacaga    120
agcgacagaag caaagcccag gcagaaccat gctaacctta cagctcagcc tgcacagaag    180
cgcagaagca aagcccaggc agaaccatgc taaccttaca gctcagcctg cacagaagcg    240
cagaagcaaa gcccaggcag aacatgctaa cettacagct cagcctgcac agaagcacag    300
a                                          301
```

<210> 283

<211> 301

<212> DNA

<213> Homo sapien

<400> 283

```
atctgtatac ggcagacaaa ctttatarag tgtagagagg tgagcgaaag gatgcaaaag    60
cactttgagg gctttataat aatatgctgc ttgaaaaaaa aaatgtgtag ttgatactca    120
gtgcatctcc agacatagta aggggttgct ctgaccaatc aggtgatcat tttttctatc    180
acttcccagg ttttatgcaa aaattttgtt aaattctata atggtgatat gcatctttta    240
ggaaacatat acatttttta aaatctatct tatgtaagaa ctgacagacg aatttgcttt    300
g                                          301
```

<210> 284

<211> 301

<212> DNA

<213> Homo sapien

<400> 284

```
caggtacaaa acgctattaa gtggcttaga atttgaacat ttgtggtctt tatttacttt    60
gcttcgtgtg tgggcaaagc aacatcttcc ctaaataatat attaccaaga aaagcaagaa    120
gcagattagg tttttgacaa aacaaacagg ccaaaagggg gctgacctgg agcagagcat    180
ggtgagaggc aaggcatgag agggcaagtt tggtgtggac agatctgtgc ctactttatt    240
actggagtaa aagaaaacaa agttcattga tgtcgaagga tatatacagt gttagaaatt    300
a                                          301
```

<210> 285

<211> 301

<212> DNA

<213> Homo sapien

<400> 285

<210> 286

<211> 301

<212> DNA

<213> Homo sapien

<400> 286

<210> 287

<211> 301

<212> DNA

<213> Homo sapien

<400> 287

<210> 288

<211> 301

<212> DNA

<213> Homo sapien

<400> 288

<210> 289

<211> 301

<222> (1)...(301)

<223> n = A,T,C or G

<400> 292

```
accttttagt agtaatgtct aataataaat aagaaatcaa ttttataagg tccatatagc      60
tgtattaaat aattttttaag tttaaaagat aaaataccat catttttaa gttgggtattc      120
aaaaccaaag natataaccg aaaggaaaaa cagatgagac ataaaatgat ttgcnagatg      180
ggaaatatag tasttyatga atgttnatta aattccagtt ataatagtgg ctacacactc      240
tcactacaca cacagacccc acagtcctat atgccacaaa cacatttcca taacttgaaa      300
a                                                                                   301
```

<210> 293

<211> 301

<212> DNA

<213> Homo sapien

<400> 293

```
ggtaccaagt gctgggtgcc gctgttacc tgttctcact gaaaagtctg gctaagtctc      60
ttgtgtagtc acttctgatt ctgacaatca atcaatcaat ggcctagagc actgactgtt      120
aacacaaacg tctactagca agtagcaaca gctttaagtc taaatacaaa gctgttctgt      180
gtgagaatth tttaaaaggc tactttgtata ataacccttg tcatthttta tgtacctcgg      240
ccgcgaccac gctaagccga attctgcaga tatccatcac actggcgggc gctcgagcat      300
g                                                                                   301
```

<210> 294

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 294

```
tgaccataa caatatacac tagctatctt tttaaactgtc catcattagc accaatgaag      60
attcaataaa attaccttta ttcacacatc tcaaaaacaat tctgcaaatt cttagtgaag      120
tttaactata gtcacaganc ttaaattatc acattgtttt ctatgtctac tgaaaataag      180
ttcactactt ttctgggata ttcttttaca aatcttatta aaattcctgg tattatcacc      240
cccaattata cagtagcaca accaccttat gtagttttta catgatagct ctgtagaggt      300
t                                                                                   301
```

<210> 295

<211> 305

<212> DNA

<213> Homo sapien

<400> 295

```
gtactctttc tctcccctcc tctgaattta attctttcaa cttgcaattt gcaaggatta      60
cacatttcac tgtgatgtat attgtgttgc aaaaaaaaaa gtgtctttgt ttaaaattac      120
ttggtttgtg aatccatctt gctttttccc cattggaact agtcattaac ccatctctga      180
actggtagaa aaacrtctga agagctagtc tatcagcatc tgacagggtga attggatggg      240
tctcagaacc atttcaccca gacagcctgt ttctatcctg tttaataaat tagtttgggt      300
tctct                                                                                   305
```

<400> 296

```
<220>
<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G
```

```
<210> 298
<211> 301
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G
```

```
<210> 299
<211> 301
<212> DNA
<213> Homo sapien
```


<400> 299

```

gttttgagac ggagttttcac tcttggttgcc cagactggac tgcaatggca gggctctctgc      60
tcactgcacc ctctgcctcc caggttcgag caattctcct gcctcagcct cccaggtagc      120
tggttgatgca ggctcacgcc accataccca gctaattttt ttgtattttt agtagagacg      180
gagtttcgcc atgttggtcca gctgggtctca aactcctgac ctcaagcgac ctgcctgcct      240
cggcctccca aagtgtctgga attataggca tgagtcaaca cgcccagcct aaagatattt      300
t

```

<210> 300

<211> 301

<212> DNA

<213> Homo sapien

<400> 300

```

attcagtttt atttgctgcc ccagtatctg taaccaggag tgccacaaaa tcttgccaga      60
tatgtccac acccactggg aaaggctccc acctggctac ttctctctatc agctgggtca      120
gctgcattcc acaagggttct cagcctaatt agtttacta cctgccagtc tcaaaactta      180
gtaaagcaag accatgacat tccccacgg aaatcagagt ttgccccacc gtcttggttac      240
tataaagcct gcctctaaca gtccttgctt cttcacacca atcccagcgc catcccccat      300
g

```

<210> 301

<211> 301

<212> DNA

<213> Homo sapien

<400> 301

```

ttaaattttt gagaggataa aaaggacaaa taatctagaa atgtgtcttc ttcagtctgc      60
agaggacccc aggtctccaa gcaaccacat ggtcaagggc atgaataatt aaaagttggt      120
gggaactcac aaagaccctc agagctgaga caccacacac agtgggagct cacaagacc      180
ctcagagctg agacaccac aacagtggga gctcacaaag accctcagag ctgagacacc      240
cacaacagca cctcgttcag ctgccacatg tgtgaataag gatgcaatgt ccagaagtgt      300
t

```

<210> 302

<211> 301

<212> DNA

<213> Homo sapien

<400> 302

```

aggtacacat ttagcttgtg gtaaatgact cacaaaaactg attttaaaat caagttaatg      60
tgaattttga aaattactac ttaatcctaa ttcacaataa caatggcatt aaggtttgac      120
ttgagttggt tcttagtatt atttatggta aataggctct taccacttgc aaataactgg      180
ccacatcatt aatgactgac ttcccagtaa ggctctctaa ggggtaagta ggaggatcca      240
caggatttga gatgctaagg ccccagagat cgtttgatcc aaccctctta ttttcagagg      300
g

```

<210> 303

<211> 301

<212> DNA

<213> Homo sapien

<400> 303

```

aggtaccaac tgtggaaata ggtagaggat cattttttct ttccatatca actaagttgt      60

```

```

atattgtttt ttgacagttt aacacatctt cttctgtcag agattctttc acaatagcac      120
tggctaattg aactaccgct tgcattgtta aaatgggtgg ttgtgaaatg atcataggcc      180
agtaacgggt atgtttttct aactgatctt ttgtctgttc caaagggacc tcaagacttc      240
catcgatttt atatctgggg tctagaaaag gagttaatct gttttccctc ataaattcac      300
c                                                                           301

```

```

<210> 304
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 304
acatggatgt tattttgcag actgtcaacc tgaatttgta tttgcttgac attgcctaatt      60
tattagtttc agtttcagct taccactttt ttgtctgcaa catgcaraas agacagtgcc      120
cttttttagtg tatcatatca ggaatcatct cacattgggt tgtgccatta ctgggtgcagt      180
gactttcagc cacttgggta aggtggagtt ggccatatgt ctccactgca aaattactga      240
ttttcctttt gtaattaata agtgtgtgtg tgaagattct ttgagatgag gtatatatct      300
c                                                                           301

```

```

<210> 305
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 305
gangtacagc gtgggtcaagg taacaagaag aaaaaaatgt gagtggcatc ctgggatgag      60
caggggggaca gacctggaca gacacgttgt catttgctgc tgtgggtagg aaaatgggcg      120
taaaggagga gaaacagata caaatctcc aactcagtat taaggatttc tcatgcctag      180
aatattggta gaaacaagaa tacattcata tggcaaataa ctaaccatgg tggaacaaaa      240
ttctgggatt taagtgggat accaangaaa ttgtattaaa agagctgttc atggaataag      300
a                                                                           301

```

```

<210> 306
<211> 8
<212> PRT
<213> Homo sapien

```

```

<400> 306
Val Leu Gly Trp Val Ala Glu Leu
1                               5

```

```

<210> 307
<211> 637
<212> DNA
<213> Homo sapien

```

```

<400> 307
acaggggratg aagggaaagg gagaggatga ggaagcccc ctggggattt ggtttggtcc      60
ttgtgatcag gtgggtctatg gggcttatcc ctacaaagaa gaatccagaa ataggggcac      120

```

```

attgaggaat gatacttgag cccaaagagc attcaatcat tgttttattt gccttmtttt 180
cacaccattg gtgagggagg gattaccacc ctgggggttat gaagatggtt gaacacccca 240
cacatagcac cggagatatg agatcaacag tttcttagcc atagagattc acagcccaga 300
gcaggaggac gcttgcacac catgcaggat gacatggggg atgcgctcgg gattgggtgtg 360
aagaagcaag gactgttaga ggcaggcttt atagtaacaa gacgggtggg caaactctga 420
tttccgtggg ggaatgtcat ggtcttgctt tactaagttt tgagactggc aggtagtga 480
actcattagg ctgagaacct tgtggaatgc acttgaccca sctgatagag gaagtagcca 540
ggtgggagcc tttcccagtg ggtgtgggac atatctggca agattttgtg gcactcctgg 600
ttacagatac tggggcagca aataaaactg aatcttg 637

```

<210> 308

<211> 647

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(647)

<223> n = A,T,C or G

<400> 308

```

acgattttca ttatcatgta aatcgggtca ctcaaggggc caaccacagc tgggagccac 60
tgctcagggg aaggttcata tgggactttc tactgccccaa gggtctatac aggatataaa 120
ggngcctcac agtatagatc tggtagcaaa gaagaagaaa caaacactga tctctttctg 180
ccacccctct gacccttttg aactcctctg acccttttaga acaagcctac ctaatatctg 240
ctagagaaaa gaccaacaac ggcctcaaag gatctcttac catgaaggtc tcagctaatt 300
cttggctaag atgtgggttc cacattaggt tctgaatatg gggggaaggg tcaatttgct 360
cattttgtgt gtggataaag tcaggatgcc cagggggcag agcagggggc tgcttgcttt 420
gggaacaatg gctgagcata taaccatagg ttatggggaa caaaacaaca tcaaagtcac 480
tgtatcaatt gccatgaaga cttgagggac ctgaatctac cgattcatct taaggcagca 540
ggaccagttt gagtggcaac aatgcagcag cagaatcaat ggaaacaaca gaatgattgc 600
aatgtccttt tttttctctt gcttctgact tgataaaagg ggaccgt 647

```

<210> 309

<211> 460

<212> DNA

<213> Homo sapien

<400> 309

```

acttttatagt ttaggctgga cattggaaaa aaaaaaaagc cagaacaaca tgtgatagat 60
aatatgattg gctgcacact tccagactga tgaatgatga acgtgatgga ctattgtatg 120
gagcacatct tcagcaagag ggggaaatac tcatcatttt tggccagcag ttgtttgatc 180
accaaacatc atgccagaat actcagcaaa cttcttagc tcttgagaag tcaaagtcg 240
ggggaattta ttcttgcaa ttttaattgg actccttatg tgagagcagc ggctaccag 300
ctgggggtgg ggagcgaacc cgtcactagt ggacatgcag tggcagagct cctggtaacc 360
acctagagga atacacaggc acatgtgtga tgccaagcgt gacacctgta gcactcaa 420
ttgtcttggt tttgtcttct ggtgtgtaag attcttaagt 460

```

<210> 310

<211> 539

<212> DNA

<213> Homo sapien

<400> 310

```
<210> 311
<211> 526
<212> DNA
<213> Homo sapien
```

<400> 311

```
<210> 312
<211> 500
<212> DNA
<213> Homo sapien
```

<400> 312

<210> 313
<211> 718

<212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(718)
 <223> n = A,T,C or G

<400> 313
 ggagatttgt gtggtttgca gccgaggag accaggaaga tctgcatggt gggaaggacc 60
 tgatgataca gaggtgagaa ataagaaagg ctgctgactt taccatctga ggccacacat 120
 ctgctgaaat ggagataatt aacatcacta gaaacagcaa gatgacaata taatgtctaa 180
 gtagtgacat gtttttgacac atttccagcc ctttttaaata tccacacaca caggaagcac 240
 aaaaggaagc acagagatcc ctgggagaaa tgcccggccg ccatcttggg tcatcgatga 300
 gcctcgccct gtgcctgntc ccgcttgtag gggaaggaca ttagaaaatg aattgatgtg 360
 ttccttaaag gatggcagga aaacagatcc tgttgtaggat atttatttga acgggattac 420
 agatttgaag tgaagtcaca aagttagcat taccaatgag aggaaaacag acgagaaaat 480
 cttgatgggt cacaagacat gcaacaaaca aaatggaata ctgtgatgac acgagcagcc 540
 aactggggag gagataccac ggggcagagg tcaggattct ggccctgctg cctaactgtg 600
 cgttatacca atcatttcta tttctaccct caaacaagct gtngaataatc tgacttacgg 660
 ttcttntggc ccacattttc atnatccacc cntcntttt aannttantc caaantgt 718

<210> 314
 <211> 358
 <212> DNA
 <213> Homo sapien

<400> 314
 gtttattttac attacagaaa aaacatcaag acaatgtata ctattttcaa tatatccata 60
 cataatcaaa tatagctgta gtacatgttt tcattgggtg agattaccac aaatgcaagg 120
 caacatgtgt agatctcttg tcttattctt ttgtctataa tactgtattg ttagtccaa 180
 gctctcggtg gtccagccac tgtgaaacat gtcctcttta gattaacctc gtggacgctc 240
 ttgttgattt gctgaactgt agtgccctgt attttgcttc tgtctgtgaa ttctgttgct 300
 tctggggcat ttccttgtag tgcagaggac caccacacag atgacagcaa tctgaatt 358

<210> 315
 <211> 341
 <212> DNA
 <213> Homo sapien

<400> 315
 taccacctcc ccgctggcac tgatgagccg catcaccatg gtcaccagca ccatgaaggc 60
 ataggtgatg atgaggacat ggaatgggcc cccaaggatg gtctgtccaa agaagcgagt 120
 gacccccatt ctgaagatgt ctggaacctc taccagcagg atgatgatag ccccaatgac 180
 agtcaccagc tccccgacca gccggatata gtccttaggg gtcatgtagg cttoctgaag 240
 tagcttctgc tgtaagaggg tgttgctccg ggggctcgtg cggttatttg tccctgggctt 300
 gaggggggcg tagatgcagc acatggtgaa gcagatgatg t 341

<210> 316
 <211> 151
 <212> DNA
 <213> Homo sapien

<400> 316

```
<210> 317
<211> 151
<212> DNA
<213> Homo sapien
```

```
<210> 318
<211> 151
<212> DNA
<213> Homo sapien
```

```
<210> 319
<211> 151
<212> DNA
<213> Homo sapien
```

```
<210> 320
<211> 150
<212> DNA
<213> Homo sapien
```

```
<210> 321
<211> 151
<212> DNA
<213> Homo sapien
```

```

      <400> 321
agcaactttg tttttcatcc aggttatttt aggottagga tttcctctca cactgcagtt      60
taggggtggca ttgtaaccag ctatggcata ggtgttaacc aaaggctgag taaacatggg      120
tgccctctgag aaatcaaagt cttcatacac t                                     151

```

<210> 322
 <211> 151
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(151)
 <223> n = A,T,C or G

<400> 322
 atccagcatc ttctcctggt tcttgccctc ctttttcttc ttcttasatt ctgcttgagg 60
 tttgggcttg gtcagtttgc cacagggctt ggagatgggt acagtcttct ggcattcggc 120
 attgtgcagg gctcgttca nacttccagt t 151

<210> 323
 <211> 151
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(151)
 <223> n = A,T,C or G

<400> 323
 tgaggacttg tktttctttt ctttattttt aatcctctta ckttgtaaatt atattgccta 60
 nagactcant tactaccag tttgtgggtt twtgggagaa atgtaactgg acagttagct 120
 gttcaatyaa aaagacactt ancccatgtg g 151

<210> 324
 <211> 461
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(461)
 <223> n = A,T,C or G

<400> 324
 acctgtgtgg aattttcagct ttctcatgc aaaaggattt tgtatccccg gcctacttga 60
 agaagtgggc agctaaagga atccaggttg ttgggtggac tgtaataacc tttgatgaaa 120
 agagttacta cgaatcccat cttggttcca gctatatcac tgacagcatg gtagaagact 180
 gcgaacctca cttctagact ttcacgggtg gacgaaacgg gttcagaaac tgccaggggc 240
 ctcatacagg gatatacaaaa taccctttgt gctaccacgg ccctggggaa tcaggtgact 300
 cacacaaaatg caatagttgg tcaactgcatt tttacctgaa ccaaagctaa acccgtgtgt 360
 gccaccatgc accatggcat gccagagttc aacactgttg ctcttgaaaa ttgggtctga 420
 aaaaacgcac aagagccctt gccctgcctt agctgangca c 461

<210> 325
 <211> 400
 <212> DNA
 <213> Homo sapien

<400> 325

acactgtttc	catgttatgt	ttctacacat	tgctacctca	gtgctcctgg	aaacttagct	60
tttgatgtct	ccaagtagtc	caccttcatt	taactctttg	aaactgtatc	atctttgcca	120
agtaagagtg	gtggcctatt	tcagctgctt	tgacaaaatg	actggctcct	gacttaacgt	180
tctataaatg	aatgtgctga	agcaaagtgc	ccatggtggc	ggcgaagaag	agaaagatgt	240
gttttgTTTT	ggactctctg	tggtcccttc	caatgctgtg	ggtttccaac	caggggaagg	300
gtcccttttg	cattgccaag	tgccataacc	atgagcacta	cgctaccatg	gttctgcctc	360
ctggccaage	aggctggttt	gcaagaatga	aatgaatgat			400

<210> 326

<211> 1215

<212> DNA

<213> Homo sapien

<400> 326

ggaggactgc	agcccgact	cgcagccctg	gcaggcggca	ctgggtcatgg	aaaacgaatt	60
gttctgctcg	ggcgtcctgg	tgcattccgca	gtgggtgctg	tcagccgcac	actgtttcca	120
gaactcctac	accatcgggc	tgggcctgca	cagtcttgag	gccgaccaag	agccagggag	180
ccagatggtg	gaggccagcc	tctccgtacg	gcacccagag	tacaacagac	ccttgctcgc	240
taacgacctc	atgctcatca	agttggacga	atccgtgtcc	gagtctgaca	ccatccggag	300
catcagcatt	gcttcgcagt	gccctaccgc	ggggaaactct	tgctcgtttt	ctggctgggg	360
tctgctggcg	aacggcagaa	tgctaccgt	gctgcagtgc	gtgaacgtgt	cgggtggtgtc	420
tgaggaggtc	tgcaagtaagc	tctatgaccc	gctgtaccac	cccagcatgt	tctgcgccgg	480
cggagggcaa	gaccagaagg	actcctgcaa	cggtgactct	ggggggcccc	tgatctgcaa	540
cgggtacttg	cagggccttg	tgtctttcgg	aaaagccccg	tgtggccaag	ttggcggtgcc	600
agggtgtctac	accaacctct	gcaaattcac	tgagtggata	gagaaaaccg	tccaggccag	660
ttaactcttg	ggactgggaa	cccatgaaat	tgacccccaa	atacatcctg	cgggaaggat	720
tcaggaatat	ctgttcccg	cccctcctcc	ctcaggccca	ggagtccagg	ccccagcccc	780
ctcctccctc	aaaccaaggg	tacagatccc	cagcccctcc	tccctcagac	ccaggagtcc	840
agacccccca	gcccctcctc	cctcagaccc	aggagtccag	cccctcctcc	ctcagaccca	900
ggagtccaga	ccccccagcc	cctcctccct	cagacccagg	ggtccaggcc	ccccacccct	960
cctccctcag	actcagaggt	ccaagcccc	aaccctcctc	tccccagacc	cagaggtcca	1020
ggtcccagcc	cctcctccct	cagacccagc	ggtccaatgc	cacctagact	ctccctgtac	1080
acagtgcccc	cttgtggcac	gtagcccaa	ccttaccagt	tggtttttca	ttttttgtcc	1140
ctttccctta	gatccagaaa	taaagtctaa	gagaagcgca	aaaaaaaaaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaa					1215

<210> 327

<211> 220

<212> PRT

<213> Homo sapien

<400> 327

Glu	Asp	Cys	Ser	Pro	His	Ser	Gln	Pro	Trp	Gln	Ala	Ala	Leu	Val	Met
1				5				10					15		
Glu	Asn	Glu	Leu	Phe	Cys	Ser	Gly	Val	Leu	Val	His	Pro	Gln	Trp	Val
			20					25					30		
Leu	Ser	Ala	Ala	His	Cys	Phe	Gln	Asn	Ser	Tyr	Thr	Ile	Gly	Leu	Gly
			35					40					45		
Leu	His	Ser	Leu	Glu	Ala	Asp	Gln	Glu	Pro	Gly	Ser	Gln	Met	Val	Glu
			50					55					60		
Ala	Ser	Leu	Ser	Val	Arg	His	Pro	Glu	Tyr	Asn	Arg	Pro	Leu	Leu	Ala
65							70								80


```

<210> 328
<211> 234
<212> DNA
<213> Homo sapien

<400> 328
cgctcgtctc tggtagctgc agccaaatca taaacggcga ggactgcagc ccgcactcgc      60
agccctggca ggcggcactg gtcatggaaa acgaattgtt ctgctcgggc gtctctgggtgc      120
atccgcagtg ggtgctgtca gccacacact gtttccagaa ctctacacc atcgggctgg      180
acctgcacag tcttgaqgcc qaccaagagc cagggagcca gatgggtggag gccca      234

```

```

<210> 330
<211> 70
<212> DNA
<213> Homo sapien

<400> 330
cccaacacaa tqcccccgatc ccataccctga ctccgcacctc aggatcgctc gtctctggta      60

```

gctgcagcca

70

<210> 331
 <211> 22
 <212> PRT
 <213> Homo sapien

<400> 331
 Gln His Asn Gly Pro Ile Pro Ser Leu Thr Pro Pro Ser Gly Ser Leu
 1 5 10 15
 Val Ser Gly Ser Cys Ser
 20

<210> 332
 <211> 2507
 <212> DNA
 <213> Homo sapien

<400> 332
 tgggtgccgct gcagccggca gagatgggtg agctcatgtt cccgctgttg ctccctccttc 60
 tgcccttcct tctgtatatg gctgcgcccc aaatcaggaa aatgctgtcc agtgggggtgt 120
 gtacatcaac tgttcagctt cctgggaaag tagttgtggt cacaggagct aatacaggta 180
 tcgggaagga gacagccaaa gagctggctc agagaggagc tcgagtatat ttagcttgcc 240
 gggatgtgga aaagggggaa ttgggtggcca aagagatcca gaccacgaca ggggaaccagc 300
 aggtgttggg gcggaaaactg gacctgtctg atactaagtc tattcgagct ttgtctaagg 360
 gcttccttagc tgagggaaaag cacctccacg ttttgatcaa caatgcagga gtgatgatgt 420
 gtcggtactc gaagacagca gatggctttg agatgcacat aggagtcaac cacttgggtc 480
 acttcctcct aacctatctg ctgctagaga aactaaaagga atcagcccca tcaaggatag 540
 taaatgtgtc ttccctcgca catcacctgg gaaggatcca cttccataac ctgcagggcg 600
 agaaattcta caatgcaggc ctggcctact gtcacagcaa gctagccaac atcctcttca 660
 cccaggaact ggcccgagga ctaaaaggct ctggcggttac gacgtattct gtacaccctg 720
 gcacagtcca atctgaactg gttcggcact catctttcat gagatggatg tgggtggcttt 780
 tctccttttt catcaagact cctcagcagg gagccagac cagcctgcac tgtgccttaa 840
 cagaaggtct tgagattcta agtgggaatc atttcagtga ctgtcatgtg gcatgggtct 900
 ctgccccagc tcgtaatgag actatagcaa ggcggctgtg ggacgtcagt tgtgacctgc 960
 tgggcctccc aatagactaa caggcagtg cagttggacc caagagaaga ctgcagcaga 1020
 ctacacagta cttcttgtca aaatgattct ccttcaagg tttcaaaacc tttagcaciaa 1080
 agagagcaaa accttcacgc cttgcctgct tgggtgtccag ttaaaactca gtgtactgcc 1140
 agattcgtct aaatgtctgt catgtccaga tttactttgc ttctgttact gccagagtta 1200
 ctagagatat cataatagga taagaagacc ctcatatgac ctgcacagct cattttcctt 1260
 ctgaaagaaa ctactaccta ggagaatcta agctatagca gggatgattt atgcaaattt 1320
 gaactagctt ctttggttcac aattcagttc ctcccaacca accagtcttc acttcaagag 1380
 ggccacactg caacctcagc ttaacatgaa taacaaagac tggctcagga gcagggttg 1440
 cccaggcatg gtggatcacc ggaggtcagt agttcaagac cagcctggcc aacatgggtga 1500
 aacccacact ctactaaaaa ttgtgtatat ctttggtgtg ctctcgttt atgtgtgcc 1560
 agggagtatt ttcacaaagt tcaaaacagc cacaataatc agagatggag caaaccagt 1620
 ccatccagtc tttatgcaaa tgaaatgctg caaagggaag cagattctgt atatgttgg 1680
 aactaccac caagagcaca tgggtagcag ggaagaagta aaaaaagaga aggagaatac 1740
 tggaagataa tgcacaaaat gaagggacta gttaaggatt aactagccct ttaaggatta 1800
 actagttaag gattaatagc aaaagayatt aaatatgcta acatagctat ggaggaattg 1860
 agggcaagca cccaggactg atgaggtctt aacaaaaacc agtgtggcaa aaaaaaaaaa 1920
 aaaaaaaaaa aaaaatccta aaacaaaca aacaaaaaaa acaattcttc attcagaaaa 1980
 attatcttag ggactgatat tggtaattat ggtcaattta ataataattt ggggcatttc 2040
 cttacattgt cttgacaaga ttaaaatgtc tgtgccaaaa ttttgtattt tatttggaga 2100

095674-09000

cttcttatca	aaagtaatgc	tgccaaagga	agtctaagga	attagtagtg	ttcccatcac	2160
ttgtttggag	tgtgctattc	taaaagattt	tgatttcctg	gaatgacaat	tatatatttaa	2220
cttttggtggg	ggaaagagtt	ataggaccac	agtcttcaact	tctgatactt	gtaaattaat	2280
cttttattgc	acttgttttg	accattaagc	tatatgttta	gaaatgggtca	ttttacggaa	2340
aaattagaaa	aattctgata	atagtgcaga	ataaatgaat	taatgtttta	cttaatttat	2400
attgaactgt	caatgacaaa	taaaaattct	ttttgattat	tttttgtttt	catttaccag	2460
aataaaaacg	taagaattaa	aagtttgatt	acaaaaaaa	aaaaaaa		2507

<210> 333

<211> 3030

<212> DNA

<213> Homo sapien

<400> 333

gcaggcgact	tgcgagctgg	gagcgattta	aaacgctttg	gattcccccg	gcctgggtgg	60
ggagagcgag	ctgggtgccc	cctagattcc	ccgccccgcg	acctcatgag	ccgacctctg	120
gctccatgga	gcccggcaat	tatgccacct	tgatggagc	caaggatata	gaaggcttgc	180
tgggagcggg	aggggggcgg	aatctggctg	ccactcccc	tctgaccagc	caccagcggg	240
cgcctacgct	gatgcctgct	gtcaactatg	cccccttgg	tctgccaggc	tggcgggagc	300
cgccaaagca	atgccaccca	tgccctgggg	tgcccagggg	gacgtcccca	gctcccgtgc	360
cttatggtta	ctttggaggc	gggtactact	cctgccgagt	gtcccggagc	tgcgtgaaac	420
cctgtgccc	ggcagccacc	ctggccgcgt	accccgcgga	gactcccacg	gccgggggag	480
agtaccccag	ycgccccact	gagtttgctt	tctatccggg	atatccggga	acctaccagc	540
ctatggccag	ttacctggac	gtgtctgtgg	tgcaactctt	gggtgtctct	ggagaaccgc	600
gacatgactc	cctgttgctt	gtggacagtt	accagtcttg	ggctctcgct	ggtggctgga	660
acagccagat	gtgttgccag	ggagaacaga	accaccagg	tcccttttgg	aaggcagcat	720
ttgcagactc	cagcgggcag	caccctcctg	acgcctgcgc	ctttcgctgc	ggccgcaaga	780
aacgcattcc	gtacagcaag	gggcagtgtg	gggagctgga	gcgggagtat	gcgggctaaca	840
agttcatcac	caaggacaag	aggcgcaaga	tctcggcagc	caccagcctc	tgggagcgcc	900
agattaccat	ctggtttcag	aaccgcgggg	tcaaagagaa	gaaggttctc	gccaagggtga	960
agaacagcgc	taccctttaa	gagatctcct	tgctgggtg	ggaggagcga	aagtgggggt	1020
gtcctgggga	gaccaggaac	ctgccaagcc	caggctgggg	ccaaggactc	tgctgagagg	1080
cccctagaga	caacaccctt	cccaggccac	tggctgctgg	actgttcctc	aggagcggcc	1140
tgggtaccca	gtatgtgcag	ggagacggaa	ccccatgtga	cagccactc	caccagggtt	1200
cccaaagaac	ctggcccagt	cataatcatt	catcctgaca	gtggcaataa	tcacgataac	1260
cagtactagc	tgccatgatc	gttagcctca	tattttctat	ctagagctct	gtagagcact	1320
ttagaaaccg	ctttcatgaa	ttgagcta	tatgaataaa	tttggaaggc	gatccctttg	1380
caggggaagc	ttctctcaga	cccccttcca	ttacacctct	cacctgggtg	acagcaggaa	1440
gactgaggag	aggggaaccg	gcagattcgt	tgtgtggctg	tgatgtccgt	ttagcatttt	1500
tctcagctga	cagctgggta	ggtggacaat	tgtagaggct	gtctcttctt	ccctccttgt	1560
ccaccccata	gggtgtaccc	actggtcttg	gaagcaccca	tccttaatac	gatgattttt	1620
ctgtcgtgtg	aaaatgaagc	cagcaggctg	cccctagtca	gtccttctct	ccagagaaaa	1680
agagatttga	gaaagtgcct	gggtaattca	ccattaattt	cctcccccaa	actctctgag	1740
tcttccctta	atatttctgg	tggttctgac	caaagcaggt	catggtttgt	tgagcatttg	1800
ggatcccagt	gaagtagatg	ttttagcctt	tgcatactta	gcccttccca	ggcacaacg	1860
gagtggcaga	gtgggtgcaa	ccctgttttc	ccagtccacg	tagacagatt	cacagtgcgg	1920
aattctggaa	gctggagaca	gacgggctct	ttgcagagcc	gggactctga	gagggacatg	1980
agggcctctg	cctctgtgtt	cattctctga	tgtcctgtac	ctgggctcag	tgcccggtgg	2040
gactcatctc	ctggccgcgc	agcaaagcca	gcgggttcgt	gctggctcct	cctgcacctt	2100
aggctggggg	tggggggcct	gccggcgcat	tctccacgat	tgagcgcaca	ggcctgaagt	2160
ctggacaacc	cgcagaaccg	aagctccgag	cagcgggtcg	gtggcgagta	gtggggtcgg	2220
tggcgagcag	ttggtggtgg	gccgcggccg	ccactacctc	gaggacattt	ccctcccggg	2280
gccagctctc	ctagaaaccc	cgcggcgggc	gccgcagcca	agtgtttatg	gcccgcggtc	2340
gggtgggata	ctagccctgt	ctcctctcct	gggaaggagt	gagggtgga	cgtgacttag	2400

acacctacaa	atctattttac	caaagaggag	cccgggactg	agggaaaagg	ccaaagagtg	2460
tgagtgcatt	cggactgggg	gttcagggga	agaggacgag	gaggaggaag	atgaggtcga	2520
tttcctgatt	taaaaaatcg	tccaagcccc	gtggtccagc	ttaaggtcct	cggttacatg	2580
cgcgcctcag	agcagggtcac	tttctgcctt	ccacgtcctc	cttcaaggaa	gccccatgtg	2640
ggtagctttc	aatatcgtag	gttcttactc	ctctgcctct	ataagctcaa	acccaccaac	2700
gacggggcaa	gtaaaccccc	tcctcgcgcg	acttcggaac	tggcgagagt	tcagcgcaga	2760
tgggcctgtg	gggagggggc	aagatagatg	agggggagcg	gcatggtgcg	gggtgacccc	2820
ttggagagag	gaaaaaggcc	acaagagggg	ctgccaccgc	cactaacgga	gatggccctg	2880
gtagagacct	ttgggggtct	ggaacctctg	gactccccat	gctctaactc	ccacactctg	2940
ctatcagaaa	cttaaacctg	aggattttct	ctgtttttca	ctcgcaataa	aytcagagca	3000
aacaaaaaaa	aaaaaaaaaa	aaaactcag				3030

<210> 334

<211> 2417

<212> DNA

<213> Homo sapien

<400> 334

ggcgcccgct	ctagagctag	tgggatcccc	cgggctgcac	gaattcggca	cgagtgaagt	60
ggagttttac	ctgtattgtt	ttaattttcaa	caagcctgag	gactagccac	aaatgtaccc	120
agttttacaaa	tgaggaaaca	ggtgcaaaaa	ggttggtacc	tgtcaaagg	cgtatgtggc	180
agagccaaga	tttgagccca	gttatgtctg	atgaacttag	cctatgctct	ttaaacttct	240
gaatgctgac	cattgaggat	atctaaactt	agatcaattg	cattttccct	ccaagactat	300
ttactttatca	atacaataat	accaccttta	ccaatctatt	gttttgatac	gagactcaaa	360
tatgccagat	atatgtaaaa	gcaacctaca	agctctctaa	tcatgctcac	ctaaaagatt	420
cccgggatct	aataggctca	aagaaacttc	ttctagaaat	ataaaaagaga	aaattggatt	480
atgcaaaaat	tcattattaa	tttttttcat	ccatccttta	attcagcaaa	catttatctg	540
ttgttgactt	tatgcagtat	ggccttttaa	ggattggggg	acagggtgaag	aacgggggtgc	600
cagaatgcat	cctcctacta	atgaggtcag	tacacatttg	catttttaaaa	tgccctgtcc	660
agctgggcat	ggtggatcat	gcctgtaatc	tcaacattgg	aaggccaagg	caggaggatt	720
gcttcagccc	aggagttcaa	gaccagcctg	ggcaacatag	aaagacccca	tctctcaatc	780
aatcaatcaa	tgccctgtct	ttgaaaataa	aactctttta	gaaagggttta	atgggcaggg	840
tgtggtagct	catgcctata	atacagcact	ttgggaggct	gaggcaggag	gatcacttta	900
gcccagaagt	tcaagaccag	cctgggcaac	aagtgcaccc	tcatctcaat	tttttaataa	960
aatgaataca	tacataagga	aagataaaaa	gaaaagttta	atgaaagaat	acagtataaa	1020
acaaatctct	tggacctaaa	agtatttttg	ttcaagccaa	atattgtgaa	tcacctctct	1080
gtgttgagga	tacagaatat	ctaagcccag	gaaactgagc	agaaaagttca	tgtactaact	1140
aatcaaccgg	aggcaaggca	aaaatgagac	taactaatca	atccgaggca	aggggcaaat	1200
tagacggaac	ctgactctgg	tctattaagc	gacaactttc	cctctgttgt	atttttcttt	1260
tattcaatgt	aaaaggataa	aaactctcta	aaactaaaaa	caatgtttgt	caggagttac	1320
aaaccatgac	caactaatta	tggggaatca	taaaatatga	ctgtatgaga	tcttgatggt	1380
ttacaaaagt	taccactgtg	taatcacttt	aaacattaat	gaacttaaaa	atgaatttac	1440
ggagatttga	atgtttcttt	cctgttgat	tagttggctc	aggctgccat	aacaaaatac	1500
cacagactgg	gaggcttaag	taacagaaat	tcattttctc	cagttctggg	ggctggaagt	1560
ccacgatcaa	ggtgcaggaa	aggcaggctt	cattctgagg	ccctctctct	ggctcacatg	1620
tggccaccct	cccactgcgt	gctcacatga	cctctttgtg	ctcctggaaa	gagggtgtgg	1680
gggacagagg	gaaagagaag	gagagggaa	tctctggtgt	ctcgtctttc	aaggacctta	1740
acctgggcca	ctttggccca	ggcactgtgg	ggtggggggt	tgtggctgct	ctgctctgag	1800
tggccaagat	aaagcaacag	aaaaatgtcc	aaagctgtgc	agcaaagaca	agccaccgaa	1860
cagggatctg	ctcatcagtg	tggggacctc	caagtcggcc	accctggagg	caagccccc	1920
cagagcccat	gcaaggtggc	agcagcagaa	gaaggggaatt	gtccctgtcc	ttggcacatt	1980
cctcaccgac	ctggtgatgc	tggacactgc	gatgaatggt	aatgtggatg	agaatatgat	2040
ggactcccag	aaaaggagac	ccagctgctc	aggtggctgc	aaatcattac	agccttcac	2100
ctggggagga	actggggggc	tggttctggg	tcagagagca	gccagtgag	ggtgagagct	2160

```
<210> 335
<211> 2984
<212> DNA
<213> Homo sapien
```

<400> 335						
atccctcctt	ccccactctc	ctttccagaa	ggcacttggg	gtcttatctg	ttggactctg	60
aaaacacttc	agggcgccct	ccaaggcttc	cccaaacc	taagcagccg	cagaagcgct	120
cccgagctgc	cttctccac	actcaggtga	tcgagttgga	gaggaagttc	agccatcaga	180
agtacctgtc	ggccccgtga	cggggccacc	tggccaagaa	cctcaagctc	acggagaccc	240
aagtgaagat	atgggtccag	aacagacgct	ataagactaa	gcgaaagcag	ctctcctcgg	300
agctgggaga	cttgagagaag	cactcctctt	tgcgggcct	gaaagaggag	gccttctccc	360
gggcctccct	ggtctccgtg	tataacagct	atccttacta	cccatacctg	tactgctggg	420
gcagctggag	cccagctttt	tggtaatgcc	agctcaggtg	acaaccatta	tgatcaaaaa	480
ctgccttccc	caggggtgtct	ctatgaaaa	cacaaagggc	caaggtcagg	gagcaagagg	540
tgtgcacacc	aaagctattg	gagattttgcg	tggaaatctc	asattcttca	ctggtgagac	600
aatgaacaa	cagagacagt	gaaagtttta	atacctaa	cattccccca	gtgcatactg	660
taggtcattt	tttttgcctc	tggctacctg	tttgaagggg	agagaggggaa	aatcaagtgg	720
tattttccag	cactttgtat	gattttggat	gagctgtaca	cccaaggatt	ctgttctgca	780
actccatcct	cctgtgtcac	tgaatatcaa	ctctgaaaga	gcaaacctaa	caggagaaag	840
gacaaccagg	atgaggatgt	caccaactga	attaaactta	agtccagaag	cctcctgttg	900
gccttggaat	atggccaagg	ctctctctgt	ccctgtaaaa	gagaggggca	aatagagagt	960
ctccaagaga	acgcccctcat	gctcagcaca	tatttgcatt	ggagggggag	atgggtggga	1020
ggagatgaaa	atatcagctt	ttcttattcc	tttttattcc	ttttaaaatg	gtatgccaac	1080
ttaagtattt	acaggggtggc	ccaaatagaa	caagatgcac	tcgctgtgat	tttaagacaa	1140
gctgtataaa	cagaactcca	ctgcaagagg	gggggcggg	ccaggagaat	ctccgcttgt	1200
ccaagacagg	ggcctaagga	gggtctccac	actgctgcta	ggggctgttg	cattttttta	1260
ttagtagaaa	gtggaaaggc	ctcttctcaa	cttttttccc	ttgggctgga	gaatttagaa	1320
tcagaagttt	cctggagttt	tcaggctatc	atatatactg	tatcctgaaa	ggcaacataa	1380
ttcttccttc	cctcctttta	aaatttttgg	ttcctttttg	cagcaattac	tcactaaagg	1440
gcttcttttt	agtcagattt	tttagtcttg	ctgcacctaa	cttagcctc	gcttatttag	1500
cccgagatct	ggtctttttt	tttttttttt	tttttccgtc	tccccaaagc	tttatctgtc	1560
ttgacttttt	aaaaaagttt	gggggcagat	tctgaattgg	ctaaaaagaca	tgcattttta	1620
aaactagcaa	ctcttatttc	tttcttttaa	aaatacatag	cattaaatcc	caaatcctat	1680
ttaaagacct	gacagcttga	gaaggctact	actgcattta	taggaccttc	tgggtggttct	1740
gctgttacgt	ttgaagtctg	acaatccttg	agaatccttg	catgcagagg	aggtaagagg	1800
tattggattt	tcacagagga	agaacacagc	gcagaatgaa	gggccaggct	tactgagctg	1860
tccagtggag	ggctcatggg	tgggacatgg	aaaagaaggc	agcctaggcc	ctggggagcc	1920
cagtccactg	agcaagcaag	ggactgagtg	agccttttgc	aggaaaaggc	taagaaaaag	1980
gaaaaccatt	ctaaaacaca	acaagaaact	gtccaaatgc	tttgggaact	gtgtttattg	2040
cctataatgg	gtcccaaaaa	tgggtaacct	agacttcaga	gagaatgagc	agagagcaaa	2100
ggagaaatct	ggctgtcctt	ccattttcat	tctgttatct	caggtgagct	ggtagagggg	2160
agacattaga	aaaaaatgaa	acaacaaaac	aattactaat	gaggtacgct	gaggcctggg	2220
agtctcttga	ctccactact	taattccgtt	tagtgagaaa	cctttcaatt	ttcttttatt	2280
agaaggccca	gcttactgtt	ggtggcaaaa	ttgccaaact	aagttaatag	aaagtggcc	2340
aaattccacc	cattttctgt	ggttttggct	ccacattgca	atgttcaatg	ccacgtgctg	2400
ctgacaccga	ccggagtact	agccagcaca	aaaggcaggg	tagcctgaat	tgctttctgc	2460
tctttacatt	tcttttaaaa	taagcattta	gtgctcagtc	cctactgaqt	actctttctc	2520

tccccctcctc tgaatttaaat tcttttcaact tgcaattttgc aaggattaca catttcactg 2580
 tgatgtatat tgtgtttgcaa aaaaaaaaaa aagtgtcttt gtttaaaatt acttggtttg 2640
 tgaatccatc ttgctttttc cccattggaa ctagtcatta acccatctct gaactggtag 2700
 aaaaacatct gaagagctag tctatcagca tctgacaggt gaattggatg gttctcagaa 2760
 ccatttcacc cagacagcct gtttctatcc tgtttaataa attagtttgg gttctctaca 2820
 tgcataacaa accctgtctc aatctgtcac ataaaagtct gtgacttgaa gtttagtcag 2880
 cccccccacc aaactttatt tttctatgtg ttttttgcaa catatgagtg ttttgaaaat 2940
 aaagtaccca tgtctttatt agaaaaaaaa aaaaaaaaaa aaaa 2984

<210> 336
 <211> 147
 <212> PRT
 <213> Homo sapien

<400> 336
 Pro Ser Phe Pro Thr Leu Leu Ser Arg Arg His Leu Gly Ser Tyr Leu
 1 5 10 15
 Leu Asp Ser Glu Asn Thr Ser Gly Ala Leu Pro Arg Leu Pro Gln Thr
 20 25 30
 Pro Lys Gln Pro Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln
 35 40 45
 Val Ile Glu Leu Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala
 50 55 60
 Pro Glu Arg Ala His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln
 65 70 75 80
 Val Lys Ile Trp Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
 85 90 95
 Leu Ser Ser Glu Leu Gly Asp Leu Glu Lys His Ser Ser Leu Pro Ala
 100 105 110
 Leu Lys Glu Glu Ala Phe Ser Arg Ala Ser Leu Val Ser Val Tyr Asn
 115 120 125
 Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr Cys Val Gly Ser Trp Ser Pro
 130 135 140
 Ala Phe Trp
 145

<210> 337
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 337
 Ala Leu Thr Gly Phe Thr Phe Ser Ala
 1 5

<210> 338
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 338
 Leu Leu Ala Asn Asp Leu Met Leu Ile
 1 5

000000"000000"000000

```
<210> 340
<211> 483
<212> DNA
<213> Homo sapien
```

<400> 340

<400> 344

```
<210> 345
<211> 251
<212> DNA
<213> Homo sapien
```

<400> 345

```
<210> 346
<211> 282
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(282)
<223> n = A,T,C or G
```

<400> 346

```
<210> 347
<211> 201
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(201)
<223> n = A,T,C or G
```

<400> 347
acacacataa tattataaaa tgccatctaa ttggaaggag ctttctatca ttgcaagtca 60
taaataataac ttttaaaaana ntactancag cttttaccta ngctcctaaa tgcttgtaaa 120
tctgagactg actggaccca cccagaccca gggcaaagat acatgttacc atatcatctt 180
tataaagaat ttttttttgt c 201

<210> 348
<211> 251
<212> DNA
<213> Homo sapien

<400> 348
ctgttaatca caacatttgt gcatcacttg tgccaagtga gaaaatgttc taaaatcaca 60
agagagaaca gtgccagaat gaaactgacc ctaagtccca ggtgcccctg ggcaggcaga 120
aggagacact cccagcatgg aggagggttt atcttttcat cctaggtcag gtctacaatg 180
ggggaagggtt ttattataga actcccaaca gccacactca ctcttgccac ccacccgatg 240
gccctgcctc c 251

<210> 349
<211> 251
<212> DNA
<213> Homo sapien

<400> 349
taaaaatcaa gccatttaat tgtatctttg aaggtaaaca atatatggga gctggatcac 60
aacccttgag gatgccagag ctatgggtcc agaacatggt gtggtattat caacagagtt 120
cagaaggggtc tgaactctac gtgttaccag agaacataat gcaattcatg cattccactt 180
agcaattttg taaaatacca gaaacagacc ccaagagtct ttcaagatga ggaaaaattca 240
actcctggtt t 251

<210> 350
<211> 908
<212> DNA
<213> Homo sapien

<400> 350
ctggacactt tgcgaggggt tttgctgggt gctgctgctg cccgtcatgc tactcatcgt 60
agcccggccg gtgaagctcg ctgctttccc tacctcctta agtgactgcc aaacgcccac 120
cggctggaat tgctctgggt atgatgacag agaaaatgat ctcttcctct gtgacaccaa 180
cacctgtaaa tttgatgggg aatgtttaag aattggagac actgtgactt gcgtctgtca 240
gttcaagtgc aacaatgact atgtgcctgt gtgtggctcc aatggggaga gctaccagaa 300
tgagtgttac ctgcgacagg ctgcatgcaa acagcagagt gagatacttg tgggtgcaga 360
aggatcatgt gccacagtcc atgaaggctc tggagaaaact agtcaaaaagg agacatccac 420
ctgtgatatt tgccagtttg gtgcagaatg tgacgaagat gccgaggatg tctggtgtgt 480
gtgtaaatatt gactgttctc aaaccaactt caatcccctc tgcgcttctg atgggaaatc 540
ttatgataat gcatgccaaa tcaaagaagc atcgtgtcag aaacaggaga aaattgaagt 600
catgtctttg ggtcgatgtc aagataacac aactacaact actaagtctg aagatgggca 660
ttatgcaaga acagattatg cagagaatgc taacaaatta gaagaaagtg ccagagaaca 720
ccacatacct tgtccggaac attacaatgg cttctgcatg catgggaagt gtgagcattc 780
tatcaatatg caggagccat cttgcagggt tgatgctggt tatactggac aacactgtga 840
aaaaaaggac tacagtgttc tatacgttgt tcccggctct gtacgatttc agtatgtctt 900
aatcgcag 908

<210> 351
 <211> 472
 <212> DNA
 <213> Homo sapien

<400> 351
 ccagttat ttt gcaagtggta agagcctatt taccataaat aataactaaga accaactcaa 60
 gtcaaacc tt aatgccattg ttattgtgaa ttaggattaa gtagtaattt tcaaaattca 120
 cattaacttg attttaaaat cagwtttgyg agtcatttac cacaagctaa atgtgtacac 180
 tatgataaaa acaaccattg tttcctgtt tttctaaaca gtccataattt ctaacactgt 240
 atatatcctt cgacatcaat gaactttgtt ttcttttact ccagtaataa agtaggcaca 300
 gatctgtcca caacaaactt gccctctcat gccttgccctc tcaccatgct ctgctccagg 360
 tcagccccct tttggcctgt ttgttttgtc aaaaacctaa tctgcttctt gctttttcttg 420
 gtaatatata tttagggaag atgttgcttt gcccacacac gaagcaaagt aa 472

<210> 352
 <211> 251
 <212> DNA
 <213> Homo sapien

<400> 352
 ctcaaagcta atctctcggg aatcaaacca gaaaagggca aggatcttag gcatgggtgga 60
 tgtggataag gccagggtcaa tggctgcaag catgcagaga aagagggtaca tcggagcgtg 120
 caggctgctg tccgtcctta cgatgaagac cacgatgcag tttccaaaca ttgccactac 180
 atacatggaa aggaggggga agccaacca gaaatgggct ttctctaata ctgggataacc 240
 aataagcaca a 251

<210> 353
 <211> 436
 <212> DNA
 <213> Homo sapien

<400> 353
 tttttttttt tttttttttt tttttttacaa caatgcagtc atttattttat tgagtatgtg 60
 cacattatgg tattattact atactgatta tattttatcat gtgacttcta attaraaaat 120
 gtatccaaaa gcaaaacagc agatatacaa aattaaagag acagaagata gacattaaca 180
 gataaggcaa cttatacatt gacaatccaa atccaatata tttaaacatt tgggaaatga 240
 gggggacaaa tggaagccar atcaaatttg tgtaaaacta ttcagtatgt ttcccttgc 300
 tcatgtctga raaggctctc ccttcaatgg ggatgacaaa ctccaaatgc cacacaaatg 360
 ttaacagaat actagattca cactggaacg ggggtaaaga agaaattatt ttctataaaa 420
 gggctcctaa tgtagt 436

<210> 354
 <211> 854
 <212> DNA
 <213> Homo sapien

<400> 354
 cctttttctag ttcaccagtt ttctgcaagg atgctgggta gggagtgtct gcaggaggag 60
 caagtctgaa accaaatcta ggaaacatag gaaacgagcc aggcacaggg ctggtgggccc 120
 atcagggaacc accctttggg ttgatatttt gcttaatctg catcttttga gtaagatcat 180
 ctggcagtag aagctgttct ccagggtacat ttctctagct catgtacaaa aacatcctga 240
 aggactttgt cagggtgcctt gctaaaagcc agatgcgttc ggcacttcct tggctctgagg 300
 ttaattgcac acctacaggc actgggctca tgctttcaag tattttgtcc tcacttttagg 360

```
<210> 355
<211> 676
<212> DNA
<213> Homo sapien
```

```
<210> 356
<211> 574
<212> DNA
<213> Homo sapien
```

```
<210> 357
<211> 393
<212> DNA
<213> Homo sapien
```

```

<400> 357
ttttttttttt tttttttttt tttttttttt tacagaatat aratgcttta tcaactgkact      60
taatatgqkg kcttqtccac tatacttaaa aatgcaccac tcataaatat ttaattcagc      120

```

aagccacaac	caaracttga	ttttatcaac	aaaaacccct	aatataaac	ggsaaaaaag	180
atagatataa	ttattccagt	ttttttaaaa	cttaaaarat	attccattgc	cgaattaara	240
araarataag	tgttatatgg	aaagaagggc	attcaagcac	actaaaraaa	cctgaggkaa	300
gcataatctg	tacaaaatta	aactgtcctt	tttggcattt	taacaaattt	gcaacgktct	360
tttttttctt	tttctgtttt	tttttttttt	tac			393

<210> 358

<211> 630

<212> DNA

<213> Homo sapien

<400> 358

acagggtaaa	caggaggatc	cttgctctca	cggagcttac	attctagcag	gaggacaata	60
ttaatgttta	taggaaaatg	atgagtttat	gacaaaggaa	gtagatagtg	ttttacaaga	120
gcatagagta	gggaagctaa	tccagcacag	ggaggtcaca	gagacatccc	taaggaagtg	180
gagtttaaac	tgagagaagc	aagtgcctaa	actgaaggat	gtgttgaaga	agaagggaga	240
gtagaacaat	ttgggcagag	ggaaccttat	agaccctaag	gtgggaagg	tcaaagaact	300
gaaagagagc	tagaacagct	ggagccgttc	tccggtgtaa	agaggagtca	aagagataag	360
attaagatg	tgaagattaa	gatcttggtg	gcattcaggg	attggcactt	ctacaagaaa	420
tcactgaagg	gagtaatgtg	acattacttt	tcacttcagg	atggccattc	taactccagg	480
gggtagactg	gactaggtaa	gactggaggc	aggtagacct	cttctaaggc	ctgcgatagt	540
gaaagacaaa	aataagtggg	gaaattcagg	ggatagtga	aatcagtagg	acttaatgag	600
caagccagag	gttcctccac	aacaaccagt				630

<210> 359

<211> 620

<212> DNA

<213> Homo sapien

<400> 359

acagcattcc	aaaatataca	tctagagact	aarrgtaaat	gctctatagt	gaagaagtaa	60
taattaaaaa	atgctactaa	tatagaaaat	ttataatcag	aaaaataaat	attcagggag	120
ctcaccagaa	gaataaagtg	ctctgccagt	tattaaagga	ttactgctgg	tgaattaaat	180
atggcattcc	ccaagggaaa	tagagagatt	cttctggatt	atgttcaata	tttatttcac	240
aggattaact	gttttaggaa	cagatataaa	gcttcgccac	ggaagagatg	gacaaagcac	300
aaagacaaca	tgatacctta	ggaagcaaca	ctaccctttc	aggcataaaa	tttggagaaa	360
tgcaacatta	tgcttcatga	ataatatgta	gaaagaagg	ctgatgaaaa	tgacatcctt	420
aatgtaagat	aactttataa	gaattctggg	tcaaataaaa	ttctttgaag	aaaacatcca	480
aatgtcattg	acttatcaaa	tactatcttg	gcatataacc	tatgaaggca	aaactaaaca	540
aacaaaaagc	tcacaccaaa	caaaaccatc	aacttatttt	gtattctata	acatacgaga	600
ctgtaaagat	gtgacagtgt					620

<210> 360

<211> 431

<212> DNA

<213> Homo sapien

<400> 360

aaaaaaaaa	agccagaaca	acatgtgata	gataatatga	ttggctgcac	acttccagac	60
tgatgaatga	tgaacgtgat	ggactattgt	atggagcaca	tcttcagcaa	gagggggaaa	120
tactcatcat	ttttggccag	cagttgtttg	atcaccaaac	atcatgccag	aatactcagc	180
aaaccttctt	agctcttgag	aagtcaaagt	ccgggggaat	ttattcctgg	caattttaat	240
tggactcctt	atgtgagagc	agcggctacc	cagctggggg	ggtggagcga	accggtcact	300
agtggacatg	cagtggcaga	gctcctggta	accacctaga	ggaatacaca	ggcacatgtg	360

tgatgccaaag cgtgacacct gtagcactca aatttgtctt gtttttgtct ttcgggtgtgt 420
agattcttag t 431

<210> 361
<211> 351
<212> DNA
<213> Homo sapien

<400> 361
acactgattt ccgatcaaaa gaatcatcat ctttaccttg acttttcagg gaattactga 60
actttcttct cagaagatag ggcacagcca ttgccttggc ctcaacttgaa gggcttgcac 120
ttgggtcctc tggctctctg ccaagtttcc cagccactcg agggagaaat atcgggaggt 180
ttgacttcct ccggggcttt cccgagggct tcaccgtgag ccctgcggcc ctgagggctg 240
caatcctgga ttcaatgtct gaaacctcgc tctctgctg ctggacttct gaggcogtca 300
ctgccactct gtctccagc tctgacagct cctcatctgt ggtcctgttg t 351

<210> 362
<211> 463
<212> DNA
<213> Homo sapien

<400> 362
acttcatcag gccataatgg gtgcctcccg tgagaatcca agcacctttg gactgcgcga 60
tgtagatgag cgggtgaag atcttgcgca tgcgcggctt cagggcgaag ttcttggcgc 120
ccccggtcac agaaatgacc aggttgggtg ttttcaggtg ccagtgtctg gtcagcagct 180
cgtaaaggat ttccgcgtcc gtgtcgcagg acagacgtat atacttccct ttcttcccca 240
gtgtctcaaa ctgaatatcc ccaaaggcgt cggtaggaaa ttcttgggtg tgtttcttgt 300
agttccattt ctcaactttg ttgactggg tgccttccat gtgctggctc tgggcattagc 360
cacacttgca cacattctcc ctgataagca cgatgggtgtg gacagggaagg aaggatttca 420
ttgagcctgc ttatggaaac tggatttgtt agcttaaata gac 463

<210> 363
<211> 653
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (653)
<223> n = A,T,C or G

<400> 363
acccccgagt nctgnctgg catactngga acgaccaacg acacacccaa gctcggcctc 60
ctcttgngga ttctgggtga catcttcatg aatggcaacc gtgccagwga ggctgtcctc 120
tgggaggcac tacgcaagat gggactgcgt cctgggggtga gacatcctct ccttgagat 180
ctaacgaaac ttctcaccta tgagttgtaa agcagaaata cctgnactac agacgagtgc 240
ccaacagcaa cccccggaa gtatgagttc ctctrgggcc tccgttccta ccatgagasc 300
tagcaagatg naagtgttga gantcattgc agaggttcag aaaagagacc cntcgtgact 360
ggcttgcaca gttcatggag gctgcagatg aggccttggg tgetctggat gctgctgcag 420
ctgaggccga agccccggct gaagcaagaa cccgcattgg aattggagat gaggctgtgt 480
ntgggccctg gagctgggat gacattgagt ttgagctgct gacctgggat gaggaaggag 540
attttgagga tcntgggtcc agaattccat ttaccttctg ggccagatac caccagaatg 600
cccgtctcag attccctcag acctttgccg gtcccattat tggctcstggt ggt 653

<210> 364
 <211> 401
 <212> DNA
 <213> Homo sapien

<400> 364
 actagaggaa agacgttaaa ccactctact accacttgtg gaactctcaa agggtaaattg 60
 acaaagccaa tgaatgactc taaaaacaat atttacattt aatggtttgt agacaataaa 120
 aaaacaagggt ggatagatct agaattgtaa cattttaaga aaaccatagc atttgacaga 180
 tgagaaagct caattataga tgcaaagtta taactaaact actatagtag taaagaaata 240
 catttcacac cttcatata aattcactat cttggcttga ggcactccat aaaatgtatc 300
 acgtgcatag taaatcttta tatttgctat ggcgttgcac tagaggactt ggactgcaac 360
 aagtggatgc gcggaaaatg aaatcttctt caatagccca g 401

<210> 365
 <211> 356
 <212> DNA
 <213> Homo sapien

<400> 365
 ccagtgtcat atttgggctt aaaatttcaa gaagggcact tcaaattggct ttgcatttgc 60
 atgtttcagt gctagagcgt aggaatagac cctggcgctc actgtgagat gttcttcagc 120
 taccagagca tcaagtctct gcagcaggtc attcttgggt aaagaaatga cttccacaaa 180
 ctctccatcc cctggcttgg gcttcggcct tgcgttttcg gcatcatctc cgttaattgg 240
 gactgtcacg atgtgtatag tacagtttga caagcctggg tccatacaga ccgctggaga 300
 acattcggca atgtcccctt tgtagccagt ttcttcttcg agctcccga gagcag 356

<210> 366
 <211> 1851
 <212> DNA
 <213> Homo sapien

<400> 366
 tcatcaccat tgccagcagc ggcaccgtta gtcaggtttt ctgggaatcc cacatgagta 60
 cttccgtgtt cttcattctt cttcaatagc cataaatctt ctagctctgg ctggctgttt 120
 tcaacttctt taagcctttg tgactcttcc tctgatgtca gctttaagtc ttgttctgga 180
 ttgctgtttt cagaagagat ttttaacatc tgtttttctt tgtagtcaga aagtaactgg 240
 caaattacat gatgatgact agaaacagca tactctctgg ccgtctttcc agatcttgag 300
 aagatacatc aacattttgc tcaagtagag ggctgactat acttgetgat ccacaacata 360
 cagcaagtat gagagcagtt cttccatata tatccagcgc atttaaattc gcttttttct 420
 tgattaaaaa tttcaccact tgetgtttt gctcatgtat accaagtagc agtgggtgtga 480
 ggccatgctt gttttttgat tgcatatcag caccgtataa gagcagtgtt ttggccatta 540
 atttatcttc attgtagaca gcatagtgtg gagtggattt tccatactca tctggaatat 600
 ttggatcagt gccatgttcc agcaacatta acgcacattc atcttctctg cattgtacgg 660
 cctttgtcag agctgtcctc tttttgttgt caaggacatt aagttgacat cgtctgtcca 720
 gcacgagttt tactacttct gaattcccat tggcagaggc cagatgtaga gcagtcctct 780
 tttgcttgct cctcttgctc acatccgtgt ccctgagcat gacgatgaga tcccttctgg 840
 ggactttacc ccaccaggca gctctgtgga gcttgtccag atcttctcca tggacgtggg 900
 acctgggac catgaaggcg ctgtcatcgt agtctcccca agcgaccacg ttgctcttgc 960
 cgctcccctg cagcagggga agcagtggca gcaccacttg cacctcttgc tcccaagcgt 1020
 cttcacagag gagtcgttgt ggtctccaga agtgcccacg ttgctcttgc cgctcccct 1080
 gtccatccag ggaggaagaa atgcaggaaa tgaaagatgc atgcacgatg gtatactcct 1140
 cagccatcaa acttctggac agcaggtcac ttccagcaag gtggagaaag ctgtccaccc 1200
 acagaggatg agatccagaa accacaatat ccattcacaa acaaacactt ttcagccaga 1260


```

aaaaccacct atgacaagcc cacagccaac ataatactaa atgggggaaaa gttagaagca 120
tttcctctga gaactgcaac aataaatata aggatgctgg atttttgtcaa atgccttttc 180
tgtgtctgtt gagatgctta tgtgactttg cttttaattc tgtttatgtg attatcacat 240
ttattgactt gcctgtgtta gaccggaaga gctgggggtgt ttctcaggag ccaccgtgtg 300
ctgcggcagc ttcgggataa cttgaggctg catcactggg gaagaaacac aytccgtgtc 360
gtggcgctga tggctgagga cagagcttca gtgtggcttc tctgcgactg gcttcttcgg 420
ggagttcttc cttcatagtt catccatatg gctccagagg aaaattatat tattttgtta 480
tggatgaaga gtattacgtt gtgcagatat actgcagtgt cttcatctct tgatgtgtga 540
ttgggtagggt tccaccatgt tgccgcagat gacatgattt cagtacctgt gtctggctga 600
aaagtgtttg tttgtgaatg gatattgtgg tttctggatc tcatcctctg tgggtggaca 660
gctttctcca ccttgctgga agtgacctgc tgtccagaag tttgatggct gaggagtata 720
ccatcgtgca tgcacttttc atttctgca tttcttctc cctggatgga cagggggagc 780
ggcaagagca acgtgggcac ttctggagac cacaacgact cctctgtgaa gacgcttggg 840
agcaagaggt gcaagtgggt ctgccactgc ttccctgct gcaggggagc ggcaagagca 900
acgtggtcgc ttggggagac tacgatgaca gcgccttcat ggatcccagg taccacgtcc 960
atggagaaga tctggacaag ctccacagag ctgcttgggt gggtaaagtc ccagaaagg 1020
atctcatcgt catgctcagg gacacggatg tgaacaagag ggacaagcaa aagaggactg 1080
ctctacatct ggctctgcc aatgggaatt cagaagtagt aaaactcgtg ctggacagac 1140
gatgtcaact taatgtcctt gacaacaaaa agaggacagc tctgacaaag gccgtacaat 1200
gccaggaaga tgaatgtgcg ttaatgttgc tggaacatgg cactgatcca aatattccag 1260
atgagtatgg aaataccact ctacactatg ctgtctacaa tgaagataaa ttaatggcca 1320
aagcaactgct cttatacggg gctgatatcg aatcaaaaaa caagcatggc ctcacaccac 1380
tgctacttgg tatacatgag caaaaacagc aagtggtgaa atttttaatc aagaaaaaag 1440
cgaatttaaa tgcgctggat agatatggaa gaactgctct catacttgct gtatgttgtg 1500
gatcagcaag tatagtcagc cctctacttg agcaaaatgt tgatgtatct tctcaagatc 1560
tggaaagacg gccagagagt atgctgtttc tagtcatcat catgtaattt gccagttact 1620
ttctgactac aaagaaaaac agatgttaaa aatctcttct gaaaacagca atccagaaca 1680
agacttaaaag ctgacatcag aggaagagtc acaaaggctt aaaggaaagtg aaaacagcca 1740
gccagaggca tggaaacttt taaatttaaa cttttggttt aatgtttttt ttttttgctt 1800
taataatatt agatagtccc aaatgaaatw acctatgaga ctaggctttg agaatcaata 1860
gattcttttt ttaagaatct tttggctagg agcgggtgtc cacgcctgta attccagcac 1920
cttgagaggc tgaggtgggc agatcacgag atcaggagat cgagaccatc ctggctaaca 1980
cggtgaaacc ccatctctac taaaaatata aaaacttagc tgggtgtggt ggcgggtgcc 2040
tgtagtccca gctactcagg argctgaggc aggagaatgg catgaaccgc ggaggtggag 2100
gttgcaagtga gccgagatcc gccactacac tccagcctgg gtgacagagc aagactctgt 2160
ctcaaaaaaa aaaaaaaaaa aaaa 2184

```

```

<210> 371
<211> 1855
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (1855)
<223> n = A,T,C or G

```

```

<400> 371
tgcaagcgc gccagtgct tgtgccacgt aactgacgc cccctgagat gtgcaagcgc 60
cacgcgcagc ttgcaagcgc ggcagcggct tggctggctt gtaacggctt gcacgcgcac 120
gccgcccccg cataaccgct agactggcct gtaacggctt gcaggcgcac gccgcagcgc 180
cgtaacggct tggctgccct gtaacggctt gcacgtgcat gctgcacgcg cgtaaacggc 240
ttggctggca tgtagccgct tggcttggct ttgcatttct tgctkggctk ggcgttgkty 300
tcttggaattg acgcttcctc cttggatkgc cgtttctctc ttggatkgac gtttcytyty 360

```

```

tcgcgttcct ttgctggact tgacctttty tctgctgggt ttggcattcc tttgggggtgg 420
gctgggtgtt ttctccgggg gggktkgccc ttcttggggg gggcgtgggk cgtccccagg 480
gggcgtgggc tttccccggg tgggtgtggg ttttcttggg gtgggggtggg ctgtgctggg 540
atccccctgc tgggggttggc agggattgac ttttttcttc aaacagattg gaaacccgga 600
gtaacntgct agttggtgaa actggttggg agacgcgac tgctgggtact actgtttctc 660
ctggctgtta aaagcagatg gtggctgagg ttgattcaat gccggctgct tcttctgtga 720
agaagccatt tggctctcagg agcaagatgg gcaagtgggt gcgcaactgct tccccctgctg 780
caggggggagc ggcaagagca acgtgggcac ttctggagac cacaacgact cctctgtgaa 840
gacgcttggg agcaagaggt gcaagtgggt ctgcccactg cttccccctgc tgcaggggag 900
cggcaagagc aacgtgkcg cttgggggaga ctacgatgac agcgccttca tggakcccag 960
gtaccacgtc crtggagaag atctggacaa gctccacaga gctgcctggg ggggtaaagt 1020
ccccagaaag gatctcatcg tcatgctcag ggacactgay gtgaacaaga rggacaagca 1080
aaagaggact gctctacatc tggcctctgc caatgggaat tcagaagtag taaaactcgt 1140
gctggacaga cgatgtcaac ttaatgtcct tgacaacaaa aagaggacag ctctgacaaa 1200
ggccgtacaa tgccaggaag atgaatgtgc gtaaatgttg ctggaacatg gcactgatcc 1260
aaatatcca gatgagtatg gaaataccac tctacactat gctgtctaca atgaagataa 1320
attaatggcc aaagcactgc tcttatacgg tgctgatatc gaatcaaaaa acaaggataa 1380
gatctactaa ttttatcttc aaaatactga aatgcattca ttttaacatt gacgtgtgta 1440
agggccagtc ttccgtatatt ggaagctcaa gcataacttg aatgaaaata ttttgaaatg 1500
acctaattat ctaagacttt attttaaata ttgttatattt caaagaagca ttagagggta 1560
cagttttttt tttttaaatg cacttctggg aaatactttt gttgaaaaca ctgaatttgt 1620
aaaaggtaat acttactatt tttcaatttt tccctcctag gatttttttc ccctaatagaa 1680
tgtaagatgg caaaatttgc cctgaaatag gttttacatg aaaactccaa gaaaagttaa 1740
acatgtttca gtgaatagag atcctgctcc tttggcaagt tcctaaaaaa cagtaataga 1800
tacgaggtga tgcgcctgct agtggcaagg tttaagatat ttctgatctc gtgcc 1855

```

```

<210> 372
<211> 1059
<212> DNA
<213> Homo sapien

```

```

<400> 372
gcaacgtggg cacttctgga gaccacaacg actcctctgt gaagacgctt gggagcaaga 60
ggtgcaagtg gtgctgcccc ctgcttcccc tgctgcaggg gagcggcaag agcaacgtgg 120
gcgcttgrgg agactmcgat gacagygcct tcatggagcc caggtaccac gtccgtggag 180
aagatctgga caagctccac agagctgccc tgggtgggta aagtccccag aaaggatctc 240
atcgatcatg tcagggacac tgaygtgaac aagarggaca agcaaaagag gactgctcta 300
catctggcct ctgccaatgg gaattcagaa gtagtaaaac tcstgctgga cagacgatgt 360
caacttaatg tctttgacaa caaaaagagg acagctctga yaaaggccgt acaatgccag 420
gaagatgaat gtgcgttaat gttgctggaa catggcactg atccaaatat tccagatgag 480
tatggaaata ccactctrca ctaygctrct tayaatgaag ataaattaat ggccaaagca 540
ctgctcttat ayggtgctga tatcgaatca aaaaacaagg tatagatcta ctaattttat 600
cttcaaaata ctgaaatgca ttcattttta cattgacgtg tgtaagggcc agtcttccgt 660
atttggaagc tcaagcataa cttgaatgaa aatattttga aatgacctaa ttatctaaga 720
ctttattttt aatattgtta ttttcaaaga agcattagag ggtacagttt ttttttttta 780
aatgcacttc tggtaaatac ttttgttgaa aacactgaat ttgtaaaagg taatacttac 840
tatttttcaa tttttccctc ctaggatttt tttcccctaa tgaaatgaag atggcaaaat 900
ttgccctgaa ataggtttta catgaaaact ccaagaaaag ttaaacatgt ttcagtgaat 960
agagatcctg ctcccttggc aagttcctaa aaaacagtaa tagatacgag gtgatgcgcc 1020
tgtcagtggc aaggtttaag atatttctga tctcgtgcc 1059

```

```

<210> 373
<211> 1155
<212> DNA

```

atggtggttg	aggttgattc	catgcgggct	gcctcttctg	tgaagaagcc	atttggctct	60
aggagcaaga	tgggcaagtg	gtgctgccgt	tgtctcccct	gctgcaggga	gagcggcaag	120
agcaacgtgg	gcacttctgg	agaccacgac	gactctgcta	tgaagacact	caggagcaag	180
atgggcaagt	ggtgccgcca	ctgcttcccc	tgctgcaggg	ggagtggcaa	gagcaacgtg	240
ggcgcttctg	gagaccacga	cgactctgct	atgaagacac	tcaggaacaa	gatgggcaag	300
tggtgctgcc	actgcttccc	ctgctgcagg	gggagcggca	agagcaaggt	gggcgcttgg	360
ggagactacg	atgacagtgc	cttcatggag	cccaggtacc	acgtccgtgg	agaagatctg	420
gacaagctcc	acagagctgc	ctgggtgggt	aaagtcccca	gaaaggatct	catcgtcatg	480
ctcagggaca	ctgacgtgaa	caagaaggac	aagcaaaaaga	ggactgctct	acatctggcc	540
tctgccaatg	ggaattcaga	agtagtaaaa	ctcctgctgg	acagacgatg	tcaacttaat	600
gtccttgaca	acaaaaagag	gacagctctg	ataaaggccg	tacaatgcc	ggaagatgaa	660
tgtgcgttaa	tgttgctgga	acatggcact	gatccaaata	ttccagatga	gtatggaaat	720
accactctgc	actacgctat	ctataatgaa	gataaattaa	tggccaaagc	actgctctta	780
tatggtgctg	atatcgaatc	aaaaaacaag	catggcctca	caccactggt	acttggtgta	840
catgagcaaa	aacagcaagt	cgtgaaattt	ttaatcaaga	aaaaagcgaa	tttaaattgca	900
ctggatagat	atggaaggac	tgctctcata	cttgctgtat	gttgtggatc	agcaagtata	960
gtcagccttc	tacttgagca	aaatattgat	gtatcttctc	aagatctatc	tggacagacg	1020
gccagagagt	atgctgtttc	tagtcatcat	catgtaattt	gccagttact	ttctgactac	1080
aaagaaaaac	agatgcataa	aatctcttct	gaaaacagca	atccagaaca	agacttaaaag	1140
ctgacatcag	aggaagagtc	acaaagggtc	aaaggcagtg	aaaatagcca	gccagagaaa	1200
atgtctcaag	aaccagaaat	aaataaggat	ggtgatagag	aggttgaaga	agaaatgaag	1260
aagcatgaaa	gtaataatgt	gggattacta	gaaaacctga	ctaattggtg	cactgctggc	1320
aatggtgata	atggattaat	tcctcaaagg	aagagcagaa	cacctgaaaa	tcagcaattt	1380
cctgacaacg	aaagtgaaga	gtatcacaga	atttgcgaat	tagtttctga	ctacaaagaa	1440

aaacagatgc	caaaataactc	ttctgaaaac	agcaaccacag	aacaagactt	aaagctgaca	1500
tcagaggaag	agtcacaaag	gcttgagggc	agtgaaaatg	gccagccaga	gctagaaaat	1560
tttatggcta	tcgaagaaat	gaagaagcac	ggaagtactc	atgtcggatt	cccagaaaac	1620
ctgactaatg	gtgccactgc	tggcaatggg	gatgatggat	taattcctcc	aaggaagagc	1680
agaacacctg	aaagccagca	atttcctgac	actgagaatg	aagagtatca	cagtgcagaa	1740
caaaatgata	ctcagaagca	attttgtgaa	gaacagaaca	ctggaatatt	acacgatgag	1800
attctgattc	atgaagaaaa	gcagatagaa	gtgggttgaaa	aaatgaattc	tgagctttct	1860
cttagttgta	agaaagaaaa	agacatcttg	catgaaaata	gtacgttgcg	ggaagaaatt	1920
gccatgctaa	gactggagct	agacacaatg	aaacatcaga	gccagctaaa	aaaaaaaaaa	1980
aaaaaaaaaa	aaaaaaaaaa					2000

<210> 375

<211> 2040

<212> DNA

<213> Homo sapien

<400> 375

atgggtggttg	aggttgattc	catgccggct	gcctcttctg	tgaagaagcc	atttgggtctc	60
aggagcaaga	tgggcaagtg	gtgctgccgt	tgttccccct	gctgcaggga	gagcgggcaag	120
agcaacgtgg	gcacttctgg	agaccacgac	gactctgcta	tgaagacact	caggagcaag	180
atgggcaagt	ggtgcccgca	ctgcttcccc	tgctgcaggg	ggagtggcaa	gagcaacgtg	240
ggcgcttctg	gagaccacga	cgactctgct	atgaagacac	tcaggaacaa	gatgggcaag	300
tggtgctgcc	actgcttccc	ctgctgcagg	gggagcggca	agagcaagggt	gggcgcttgg	360
ggagactacg	atgacagtgc	cttcattggag	cccaggtacc	acgtccgtgg	agaagatctg	420
gacaagctcc	acagagctgc	ctgggtgggg	aaagtcccca	gaaaggatct	catcgctcatg	480
ctcagggaca	ctgacgtgaa	caagaaggac	aagcaaaaaga	ggactgctct	acatctggcc	540
tctgccaatg	ggaattcaga	agtagtaaaa	ctctgctggt	acagacgatg	tcaacttaat	600
gtccttgaca	acaaaaagag	gacagctctg	ataaaggccg	tacaatgcc	ggaagatgaa	660
tgtgcgttaa	tggtgctgga	acatggcact	gatccaaata	ttccagatga	gtatggaaat	720
accactctgc	actacgctat	ctataatgaa	gataaattaa	tggccaaagc	actgctctta	780
tatgggtgctg	atatcgaatc	aaaaaacaaag	catggcctca	caccactggt	acttgggtgta	840
catgagcaaa	aacagcaagt	cgtgaaatct	ttaatcaaga	aaaaagcgaa	tttaaatgca	900
ctggatagat	atggaaggac	tgctctcata	cttgctgtat	gttgtggatc	agcaagtata	960
gtcagccttc	tacttgagca	aaatattgat	gtatcttctc	aagatctatc	tggacagacg	1020
gccagagagt	atgctgtttc	tagtcatcat	catgtaattt	gccagttact	ttctgactac	1080
aaagaaaaac	agatgctaaa	aatctcttct	gaaaacagca	atccagaaca	agacttaaaag	1140
ctgacatcag	aggaagagtc	acaaagggtc	aaaggcagtg	aaaatagcca	gccagagaaa	1200
atgtctcaag	aaccagaaat	aaataaggat	ggtgatagag	aggttgaaga	agaaatgaag	1260
aagcatgaaa	gtaataatgt	gggattacta	gaaaacctga	ctaattggtgt	cactgctggc	1320
aatgggtgata	atggattaat	tcctcaaagg	aagagcagaa	cacctgaaaa	tcagcaattt	1380
cctgacaacg	aaagtgaaga	gtatcacaga	atgtgcgaat	tagtttctga	ctacaaagaa	1440
aaacagatgc	caaaataactc	ttctgaaaac	agcaaccacag	aacaagactt	aaagctgaca	1500
tcagaggaag	agtcacaaag	gcttgagggc	agtgaaaatg	gccagccaga	gaaaagatct	1560
caagaaccag	aaataaataa	ggatgggtgat	agagagctag	aaaattttat	ggctatcgaa	1620
gaaatgaaga	agcacggaag	tactcatgtc	ggattcccag	aaaacctgac	taatggtgcc	1680
actgctggca	atgggtgatga	tggattaatt	cctccaagga	agagcagaac	acctgaaagc	1740
cagcaatttc	ctgacactga	gaatgaagag	tatcacagtg	acgaacaaaa	tgatactcag	1800
aagcaatttt	gtgaagaaca	gaacactgga	atattacacg	atgagattct	gattcatgaa	1860
gaaaagcaga	tagaagtggg	tgaaaaaatg	aattctgagc	tttctcttag	ttgtaagaaa	1920
gaaaagacac	tcttgcatga	aaatagtacg	ttgcgggaag	aaattgccat	gctaagactg	1980
gagctagaca	caatgaaaca	tcagagccag	ctaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2040

<210> 376

<211> 329

<213> Homo sapien

<400> 376

Met 1	Asp	Ile	Val	Val 5	Ser	Gly	Ser	His	Pro 10	Leu	Trp	Val	Asp	Ser 15	Phe
Leu	His	Leu	Ala 20	Gly	Ser	Asp	Leu	Leu 25	Ser	Arg	Ser	Leu 30	Met	Ala	Glu
Glu	Tyr	Thr	Ile 35	Val	His	Ala	Ser 40	Phe	Ile	Ser	Cys	Ile 45	Ser	Ser	Ser
Leu	Asp 50	Gly	Gln	Gly	Glu	Arg 55	Gln	Glu	Gln	Arg	Gly	His 60	Phe	Trp	Arg
Pro 65	Gln	Arg	Leu	Leu	Cys 70	Glu	Asp	Ala	Trp	Glu	Gln	Glu 75	Val	Gln	Val 80
Val	Leu	Pro	Leu 85	Leu	Pro	Leu	Leu	Gln 90	Gly	Ser	Gly	Lys 95	Ser	Asn	Val
Val	Ala	Trp	Gly 100	Asp	Tyr	Asp	Asp	Ser 105	Ala	Phe	Met	Asp 110	Pro	Arg	Tyr
His	Val	His	Gly 115	Glu	Asp	Leu	Asp 120	Lys	Leu	His	Arg	Ala 125	Ala	Trp	Trp
Gly	Lys 130	Val	Pro	Arg	Lys	Asp 135	Leu	Ile	Val	Met	Leu	Arg 140	Asp	Thr	Asp
Val 145	Asn	Lys	Arg	Asp	Lys 150	Gln	Lys	Arg	Thr	Ala	Leu	His 155	Leu	Ala	Ser 160
Ala	Asn	Gly	Asn 165	Ser	Glu	Val	Val	Lys 170	Leu	Val	Leu	Asp 175	Arg	Arg	Cys
Gln	Leu	Asn	Val 180	Leu	Asp	Asn	Lys 185	Lys	Arg	Thr	Ala	Leu 190	Thr	Lys	Ala
Val	Gln 195	Cys	Gln	Glu	Asp	Glu 200	Cys	Ala	Leu	Met	Leu	Leu 205	Glu	His	Gly
Thr	Asp 210	Pro	Asn	Ile	Pro	Asp 215	Glu	Tyr	Gly	Asn	Thr	Thr 220	Leu	His	Tyr
Ala 225	Val	Tyr	Asn 230	Glu	Asp	Lys	Leu	Met	Ala	Lys	Ala	Leu 235	Leu	Leu	Tyr 240
Gly	Ala	Asp	Ile 245	Glu	Ser	Lys	Asn	Lys	His 250	Gly	Leu	Thr 255	Pro	Leu	Leu
Leu	Gly	Ile	His 260	Glu	Gln	Lys	Gln	Gln	Val 265	Val	Lys	Phe 270	Leu	Ile	Lys
Lys	Lys 275	Ala	Asn	Leu	Asn	Ala	Leu 280	Asp	Arg	Tyr	Gly	Arg 285	Thr	Ala	Leu
Ile	Leu 290	Ala	Val	Cys	Cys	Gly	Ser 295	Ala	Ser	Ile	Val	Ser 300	Pro	Leu	Leu
Glu 305	Gln	Asn	Val	Asp	Val	Ser 310	Ser	Gln	Asp	Leu	Glu	Arg 315	Arg	Pro	Glu 320
Ser	Met	Leu	Phe 325	Leu	Val	Ile	Ile	Met							

<210> 377

<211> 148

<212> PRT

<213> Homo sapien

<220>

<221> VARIANT

<222> (1)...(148)

<223> Xaa = Any Amino Acid

<400> 377

Met	Thr	Xaa	Pro	Ser	Trp	Ser	Pro	Gly	Thr	Thr	Ser	Val	Glu	Lys	Ile
1				5					10					15	
Trp	Thr	Ser	Ser	Thr	Glu	Leu	Pro	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys
			20					25					30		
Asp	Leu	Ile	Val	Met	Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Xaa	Asp	Lys
		35					40					45			
Gln	Lys	Arg	Thr	Ala	Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu
	50					55					60				
Val	Val	Lys	Leu	Xaa	Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn	Val	Leu	Asp
65					70					75					80
Asn	Lys	Lys	Arg	Thr	Ala	Leu	Xaa	Lys	Ala	Val	Gln	Cys	Gln	Glu	Asp
				85					90					95	
Glu	Cys	Ala	Leu	Met	Leu	Leu	Glu	His	Gly	Thr	Asp	Pro	Asn	Ile	Pro
			100					105					110		
Asp	Glu	Tyr	Gly	Asn	Thr	Thr	Leu	His	Tyr	Ala	Xaa	Tyr	Asn	Glu	Asp
		115					120					125			
Lys	Leu	Met	Ala	Lys	Ala	Leu	Leu	Leu	Tyr	Gly	Ala	Asp	Ile	Glu	Ser
	130					135					140				
Lys	Asn	Lys	Val												
145															

<210> 378

<211> 1719

<212> PRT

<213> Homo sapien

<400> 378

Met	Val	Val	Glu	Val	Asp	Ser	Met	Pro	Ala	Ala	Ser	Ser	Val	Lys	Lys
1				5					10					15	
Pro	Phe	Gly	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys	Cys	Arg	Cys	Phe
			20					25					30		
Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys	Ser	Asn	Val	Gly	Thr	Ser	Gly	Asp
		35				40						45			
His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp
	50					55					60				
Cys	Arg	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly	Lys	Ser	Asn	Val
65					70				75						80
Gly	Ala	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Asn
				85				90					95		
Lys	Met	Gly	Lys	Trp	Cys	Cys	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser
			100					105					110		
Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp	Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe
		115				120					125				
Met	Glu	Pro	Arg	Tyr	His	Val	Arg	Gly	Glu	Asp	Leu	Asp	Lys	Leu	His
	130					135					140				
Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	Asp	Leu	Ile	Val	Met
145					150					155					160
Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Lys	Asp	Lys	Gln	Lys	Arg	Thr	Ala
				165					170					175	
Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	Val	Val	Lys	Leu	Leu

000000"000000

610					615					620					
Glu 625	Tyr	Gly	Asn	Thr	Thr 630	Leu	His	Tyr	Ala	Ile 635	Tyr	Asn	Glu	Asp	Lys 640
Leu	Met	Ala	Lys	Ala 645	Leu	Leu	Leu	Tyr	Gly 650	Ala	Asp	Ile	Glu	Ser	Lys 655
Asn	Lys	His	Gly 660	Leu	Thr	Pro	Leu	Leu	Leu 665	Gly	Val	His	Glu	Gln	Lys 670
Gln	Gln	Val 675	Val	Lys	Phe	Leu	Ile	Lys 680	Lys	Lys	Ala	Asn	Leu	Asn	Ala 685
Leu	Asp 690	Arg	Tyr	Gly	Arg	Thr 695	Ala	Leu	Ile	Leu	Ala 700	Val	Cys	Cys	Gly 705
Ser 705	Ala	Ser	Ile	Val	Ser 710	Leu	Leu	Leu	Glu	Gln 715	Asn	Ile	Asp	Val	Ser 720
Ser	Gln	Asp	Leu	Ser 725	Gly	Gln	Thr	Ala	Arg 730	Glu	Tyr	Ala	Val	Ser	Ser 735
His	His	His 740	Val	Ile	Cys	Gln	Leu	Leu 745	Ser	Asp	Tyr	Lys	Glu	Lys	Gln 750
Met	Leu	Lys 755	Ile	Ser	Ser	Glu	Asn	Ser 760	Asn	Pro	Glu	Gln	Asp	Leu	Lys 765
Leu	Thr 770	Ser	Glu	Glu	Glu 775	Ser	Gln	Arg	Phe	Lys	Gly 780	Ser	Glu	Asn	Ser 785
Gln 785	Pro	Glu	Lys	Met	Ser 790	Gln	Glu	Pro	Glu	Ile 795	Asn	Lys	Asp	Gly	Asp 800
Arg	Glu	Val	Glu	Glu 805	Glu	Met	Lys	Lys	His 810	Glu	Ser	Asn	Asn	Val	Gly 815
Leu	Leu	Glu 820	Asn	Leu	Thr	Asn	Gly 825	Val	Thr	Ala	Gly	Asn	Gly	Asp	Asn 830
Gly	Leu	Ile 835	Pro	Gln	Arg	Lys	Ser 840	Arg	Thr	Pro	Glu	Asn	Gln	Gln	Phe 845
Pro	Asp 850	Asn	Glu	Ser	Glu	Glu 855	Tyr	His	Arg	Ile	Cys 860	Glu	Leu	Val	Ser 865
Asp 865	Tyr	Lys	Glu	Lys	Gln 870	Met	Pro	Lys	Tyr	Ser	Ser 875	Glu	Asn	Ser	Asn 880
Pro	Glu	Gln	Asp	Leu 885	Lys	Leu	Thr	Ser	Glu 890	Glu	Ser	Gln	Arg	Leu	895
Glu	Gly	Ser 900	Glu	Asn	Gly	Gln	Pro	Glu 905	Leu	Glu	Asn	Phe	Met	Ala	Ile 910
Glu	Glu	Met 915	Lys	Lys	His	Gly	Ser 920	Thr	His	Val	Gly	Phe	Pro	Glu	Asn 925
Leu	Thr 930	Asn	Gly	Ala	Thr	Ala 935	Gly	Asn	Gly	Asp	Asp 940	Gly	Leu	Ile	Pro 945
Pro	Arg 945	Lys	Ser	Arg	Thr	Pro 950	Glu	Ser	Gln	Gln	Phe	Pro	Asp	Thr	Glu 960
Asn	Glu	Glu	Tyr	His 965	Ser	Asp	Glu	Gln	Asn 970	Asp	Thr	Gln	Lys	Gln	Phe 975
Cys	Glu	Glu 980	Gln	Asn	Thr	Gly	Ile	Leu 985	His	Asp	Glu	Ile	Leu	Ile	His 990
Glu	Glu	Lys 995	Gln	Ile	Glu	Val	Val 1000	Glu	Lys	Met	Asn	Ser	Glu	Leu	Ser 1005
Leu	Ser 1010	Cys	Lys	Lys	Glu	Lys 1015	Asp	Ile	Leu	His	Glu 1020	Asn	Ser	Thr	Leu 1025
Arg	Glu	Glu	Ile	Ala	Met 1030	Leu	Arg	Leu	Glu	Leu 1035	Asp	Thr	Met	Lys	His 1040
Gln	Ser	Gln	Leu	Pro	Arg	Thr	His	Met	Val	Val	Glu	Val	Asp	Ser	Met

1045 1050 1055
 Pro Ala Ala Ser Ser Val Lys Lys Pro Phe Gly Leu Arg Ser Lys Met
 1060 1065 1070
 Gly Lys Trp Cys Cys Arg Cys Phe Pro Cys Cys Arg Glu Ser Gly Lys
 1075 1080 1085
 Ser Asn Val Gly Thr Ser Gly Asp His Asp Asp Ser Ala Met Lys Thr
 1090 1095 1100
 Leu Arg Ser Lys Met Gly Lys Trp Cys Arg His Cys Phe Pro Cys Cys
 1105 1110 1115 1120
 Arg Gly Ser Gly Lys Ser Asn Val Gly Ala Ser Gly Asp His Asp Asp
 1125 1130 1135
 Ser Ala Met Lys Thr Leu Arg Asn Lys Met Gly Lys Trp Cys Cys His
 1140 1145 1150
 Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Lys Val Gly Ala Trp
 1155 1160 1165
 Gly Asp Tyr Asp Asp Ser Ala Phe Met Glu Pro Arg Tyr His Val Arg
 1170 1175 1180
 Gly Glu Asp Leu Asp Lys Leu His Arg Ala Ala Trp Trp Gly Lys Val
 1185 1190 1195 1200
 Pro Arg Lys Asp Leu Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys
 1205 1210 1215
 Lys Asp Lys Gln Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly
 1220 1225 1230
 Asn Ser Glu Val Val Lys Leu Leu Leu Asp Arg Arg Cys Gln Leu Asn
 1235 1240 1245
 Val Leu Asp Asn Lys Lys Arg Thr Ala Leu Ile Lys Ala Val Gln Cys
 1250 1255 1260
 Gln Glu Asp Glu Cys Ala Leu Met Leu Leu Glu His Gly Thr Asp Pro
 1265 1270 1275 1280
 Asn Ile Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr Ala Ile Tyr
 1285 1290 1295
 Asn Glu Asp Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr Gly Ala Asp
 1300 1305 1310
 Ile Glu Ser Lys Asn Lys His Gly Leu Thr Pro Leu Leu Leu Gly Val
 1315 1320 1325
 His Glu Gln Lys Gln Gln Val Val Lys Phe Leu Ile Lys Lys Lys Ala
 1330 1335 1340
 Asn Leu Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu Ile Leu Ala
 1345 1350 1355 1360
 Val Cys Cys Gly Ser Ala Ser Ile Val Ser Leu Leu Leu Glu Gln Asn
 1365 1370 1375
 Ile Asp Val Ser Ser Gln Asp Leu Ser Gly Gln Thr Ala Arg Glu Tyr
 1380 1385 1390
 Ala Val Ser Ser His His His Val Ile Cys Gln Leu Leu Ser Asp Tyr
 1395 1400 1405
 Lys Glu Lys Gln Met Leu Lys Ile Ser Ser Glu Asn Ser Asn Pro Glu
 1410 1415 1420
 Gln Asp Leu Lys Leu Thr Ser Glu Glu Glu Ser Gln Arg Phe Lys Gly
 1425 1430 1435 1440
 Ser Glu Asn Ser Gln Pro Glu Lys Met Ser Gln Glu Pro Glu Ile Asn
 1445 1450 1455
 Lys Asp Gly Asp Arg Glu Val Glu Glu Glu Met Lys Lys His Glu Ser
 1460 1465 1470
 Asn Asn Val Gly Leu Leu Glu Asn Leu Thr Asn Gly Val Thr Ala Gly

009060"6225960

1475 1480 1485
 Asn Gly Asp Asn Gly Leu Ile Pro Gln Arg Lys Ser Arg Thr Pro Glu
 1490 1495 1500
 Asn Gln Gln Phe Pro Asp Asn Glu Ser Glu Glu Tyr His Arg Ile Cys
 1505 1510 1515 1520
 Glu Leu Val Ser Asp Tyr Lys Glu Lys Gln Met Pro Lys Tyr Ser Ser
 1525 1530 1535
 Glu Asn Ser Asn Pro Glu Gln Asp Leu Lys Leu Thr Ser Glu Glu Glu
 1540 1545 1550
 Ser Gln Arg Leu Glu Gly Ser Glu Asn Gly Gln Pro Glu Lys Arg Ser
 1555 1560 1565
 Gln Glu Pro Glu Ile Asn Lys Asp Gly Asp Arg Glu Leu Glu Asn Phe
 1570 1575 1580
 Met Ala Ile Glu Glu Met Lys Lys His Gly Ser Thr His Val Gly Phe
 1585 1590 1595 1600
 Pro Glu Asn Leu Thr Asn Gly Ala Thr Ala Gly Asn Gly Asp Asp Gly
 1605 1610 1615
 Leu Ile Pro Pro Arg Lys Ser Arg Thr Pro Glu Ser Gln Gln Phe Pro
 1620 1625 1630
 Asp Thr Glu Asn Glu Glu Tyr His Ser Asp Glu Gln Asn Asp Thr Gln
 1635 1640 1645
 Lys Gln Phe Cys Glu Glu Gln Asn Thr Gly Ile Leu His Asp Glu Ile
 1650 1655 1660
 Leu Ile His Glu Glu Lys Gln Ile Glu Val Val Glu Lys Met Asn Ser
 1665 1670 1675 1680
 Glu Leu Ser Leu Ser Cys Lys Lys Glu Lys Asp Ile Leu His Glu Asn
 1685 1690 1695
 Ser Thr Leu Arg Glu Glu Ile Ala Met Leu Arg Leu Glu Leu Asp Thr
 1700 1705 1710
 Met Lys His Gln Ser Gln Leu
 1715

<210> 379
 <211> 656
 <212> PRT
 <213> Homo sapien

<400> 379
 Met Val Val Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys Lys
 1 5 10 15
 Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys Arg Cys Phe
 20 25 30
 Pro Cys Cys Arg Glu Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp
 35 40 45
 His Asp Asp Ser Ala Met Lys Thr Leu Arg Ser Lys Met Gly Lys Trp
 50 55 60
 Cys Arg His Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val
 65 70 75 80
 Gly Ala Ser Gly Asp His Asp Asp Ser Ala Met Lys Thr Leu Arg Asn
 85 90 95
 Lys Met Gly Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly Ser
 100 105 110
 Gly Lys Ser Lys Val Gly Ala Trp Gly Asp Tyr Asp Asp Ser Ala Phe
 115 120 125

Met Glu Pro Arg Tyr His Val Arg Gly Glu Asp Leu Asp Lys Leu His
 130 135 140
 Arg Ala Ala Trp Trp Gly Lys Val Pro Arg Lys Asp Leu Ile Val Met
 145 150 155 160
 Leu Arg Asp Thr Asp Val Asn Lys Lys Asp Lys Gln Lys Arg Thr Ala
 165 170 175
 Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu Val Val Lys Leu Leu
 180 185 190
 Leu Asp Arg Arg Cys Gln Leu Asn Val Leu Asp Asn Lys Lys Arg Thr
 195 200 205
 Ala Leu Ile Lys Ala Val Gln Cys Gln Glu Asp Glu Cys Ala Leu Met
 210 215 220
 Leu Leu Glu His Gly Thr Asp Pro Asn Ile Pro Asp Glu Tyr Gly Asn
 225 230 235 240
 Thr Thr Leu His Tyr Ala Ile Tyr Asn Glu Asp Lys Leu Met Ala Lys
 245 250 255
 Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu Ser Lys Asn Lys His Gly
 260 265 270
 Leu Thr Pro Leu Leu Leu Gly Val His Glu Gln Lys Gln Gln Val Val
 275 280 285
 Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr
 290 295 300
 Gly Arg Thr Ala Leu Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile
 305 310 315 320
 Val Ser Leu Leu Leu Glu Gln Asn Ile Asp Val Ser Ser Gln Asp Leu
 325 330 335
 Ser Gly Gln Thr Ala Arg Glu Tyr Ala Val Ser Ser His His His Val
 340 345 350
 Ile Cys Gln Leu Leu Ser Asp Tyr Lys Glu Lys Gln Met Leu Lys Ile
 355 360 365
 Ser Ser Glu Asn Ser Asn Pro Glu Gln Asp Leu Lys Leu Thr Ser Glu
 370 375 380
 Glu Glu Ser Gln Arg Phe Lys Gly Ser Glu Asn Ser Gln Pro Glu Lys
 385 390 395 400
 Met Ser Gln Glu Pro Glu Ile Asn Lys Asp Gly Asp Arg Glu Val Glu
 405 410 415
 Glu Glu Met Lys Lys His Glu Ser Asn Asn Val Gly Leu Leu Glu Asn
 420 425 430
 Leu Thr Asn Gly Val Thr Ala Gly Asn Gly Asp Asn Gly Leu Ile Pro
 435 440 445
 Gln Arg Lys Ser Arg Thr Pro Glu Asn Gln Gln Phe Pro Asp Asn Glu
 450 455 460
 Ser Glu Glu Tyr His Arg Ile Cys Glu Leu Val Ser Asp Tyr Lys Glu
 465 470 475 480
 Lys Gln Met Pro Lys Tyr Ser Ser Glu Asn Ser Asn Pro Glu Gln Asp
 485 490 495
 Leu Lys Leu Thr Ser Glu Glu Glu Ser Gln Arg Leu Glu Gly Ser Glu
 500 505 510
 Asn Gly Gln Pro Glu Leu Glu Asn Phe Met Ala Ile Glu Glu Met Lys
 515 520 525
 Lys His Gly Ser Thr His Val Gly Phe Pro Glu Asn Leu Thr Asn Gly
 530 535 540
 Ala Thr Ala Gly Asn Gly Asp Asp Gly Leu Ile Pro Pro Arg Lys Ser
 545 550 555 560

000000"022500

	<210> 380														
	<211> 671														
	<212> PRT														
	<213> Homo sapien														
	<400> 380														
Met	Val	Val	Glu	Val	Asp	Ser	Met	Pro	Ala	Ala	Ser	Ser	Val	Lys	Lys
1				5					10					15	
Pro	Phe	Gly	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys	Cys	Arg	Cys	Phe
			20					25					30		
Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys	Ser	Asn	Val	Gly	Thr	Ser	Gly	Asp
		35					40					45			
His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp
	50					55					60				
Cys	Arg	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly	Lys	Ser	Asn	Val
65					70					75				80	
Gly	Ala	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Asn
				85					90					95	
Lys	Met	Gly	Lys	Trp	Cys	Cys	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser
			100					105					110		
Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp	Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe
		115					120					125			
Met	Glu	Pro	Arg	Tyr	His	Val	Arg	Gly	Glu	Asp	Leu	Asp	Lys	Leu	His
	130					135					140				
Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	Asp	Leu	Ile	Val	Met
145					150					155					160
Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Lys	Asp	Lys	Gln	Lys	Arg	Thr	Ala
				165					170					175	
Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	Val	Val	Lys	Leu	Leu
			180				185						190		
Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn	Val	Leu	Asp	Asn	Lys	Lys	Arg	Thr
		195					200					205			
Ala	Leu	Ile	Lys	Ala	Val	Gln	Cys	Gln	Glu	Asp	Glu	Cys	Ala	Leu	Met
	210					215					220				
Leu	Leu	Glu	His	Gly	Thr	Asp	Pro	Asn	Ile	Pro	Asp	Glu	Tyr	Gly	Asn
225					230					235				240	
Thr	Thr	Leu	His	Tyr	Ala	Ile	Tyr	Asn	Glu	Asp	Lys	Leu	Met	Ala	Lys
				245					250					255	
Ala	Leu	Leu	Leu	Tyr	Gly	Ala	Asp	Ile	Glu	Ser	Lys	Asn	Lys	His	Gly
		260						265					270		
Leu	Thr	Pro	Leu	Leu	Leu	Gly	Val	His	Glu	Gln	Lys	Gln	Gln	Val	Val

```
<210> 380
<211> 671
<212> PRT
<213> Homo sapien
```

```
<210> 381
<211> 251
<212> DNA
<213> Homo sapien
```

```
<210> 382
<211> 3279
<212> DNA
<213> Homo sapiens
```

<400>	382					
cttcctgcag	cccccatgct	ggtgagggggc	acggggcagga	acagtggacc	caacatggaa	60
atgctggagg	gtgtcaggaa	gtgatcgggc	tctggggcag	ggaggagggg	tggggagtgt	120
cactgggagg	ggacatcctg	cagaaggtag	gagtgcagca	acaccgcgtg	caggggaggg	180
gagagccctg	cggcacctgg	gggagcagag	ggagcagcac	ctgcccaggc	ctgggaggag	240
gggcctggag	ggcgtgagga	ggagcgaggg	ggctgcatgg	ctggagtgg	ggatcagggg	300
cagggcgcg	gatggccctc	cacaggggaag	agagggcccc	tcctgcaggg	cctcacctgg	360
gccacaggag	gacactgctt	ttcctctgag	gagtcaggag	ctgtggatgg	tgctggacag	420
aagaaggaca	gggcctggct	caggtgtcca	gaggctgtcg	ctggcttccc	tttgggatca	480
gactgcaggg	agggaggggc	gcagggttgt	ggggggagtg	acgatgagga	tgacctgggg	540
gtggctccag	gcccctggccc	tgccttggccc	ctcaccagc	ctccctcaca	gtctcctggc	600
cctcagtcct	tcccctccac	tcctatccct	atctggcctc	agtgggtcat	tctgatcact	660
gaactgacca	taccagccc	tgcccacggc	cctccatggc	tcccaatgc	cctggagagg	720
ggacatctag	tcagagagta	gtcctgaaga	ggtggcctct	gcgatgtgcc	tgtgggggca	780
gcacctcgca	gatggtcccg	gccctcatcc	tgtgcacctg	tctgcaggga	ctgtcctcct	840
ggaccttgcc	ccttgtgcag	gagctggacc	ctgaagtccc	ctcccatag	gccaagactg	900
gagccttggt	cctctgtttg	gactccctgc	ccatattctt	gtgggagtg	gttctggaga	960
catttctgtc	tgttcctgag	agctgggaat	tgtctcagt	catctgcctg	cgcggttctg	1020
agagatggag	ttgcctaggc	agttattggg	gccaatcttt	ctcactgtgt	ctctcctcct	1080
ttacccttag	ggtgattctg	ggggtccact	tgtctgtaat	ggtgtgcttc	aaggatatcac	1140
atcatggggc	cctgagccat	gtgccctgcc	tgaaaagcct	gctgtgtaca	ccaagggtgt	1200
gcattaccgg	aagtggatca	aggacaccat	cgcagccaac	ccctgagtgc	ccctgtccca	1260
ccctacctc	tagtaaattt	aagtccacct	cacgttctgg	catcacttgg	cctttctgga	1320
tgctggacac	ctgaagcttg	gaactcacct	ggccgaagct	cgagcctcct	gagtcctact	1380
gacctgtgct	ttctgggttg	gagtcacagg	ctgctaggaa	aaggaattgg	cagacacagg	1440
tgtatgccaa	tgtttctgaa	atgggtataa	tttcgtctc	tccttcggaa	cactggctgt	1500
ctctgaagac	ttctcgctca	gtttcagtg	ggacacacac	aaagacgtgg	gtgaccatgt	1560
tgtttgtggg	gtgcagagat	gggaggggtg	gggcccaccc	tgggaagagt	gacagtgaca	1620
caagggtggc	actctctaca	gatcactgag	gataagctgg	agccacaatg	catgaggcac	1680
acacacagca	aggttgacgc	tgtaaacata	gcccacgctg	tcctgggggg	actgggaagc	1740
ctagataaag	ccgtgagcag	aaagaagggg	aggatcctcc	tatgttgttg	aaggagggac	1800
tagggggaga	aactgaaagc	tgattaatta	caggagggtt	gttcagggtc	cccaaaccac	1860
cgtcagatct	gatgatcttc	tagcaggact	tacagaaata	aagagctatc	atgctgtggt	1920
ttattatggt	ttgttacatt	gataggatac	atactgaaat	cagcaaaca	aacagatgta	1980
tagattagag	tgtggagaaa	acagaggaaa	acttgcagtt	acgaagactg	gcaacttggc	2040
ttactaagt	tttcagactg	gcagggaagtc	aaacctatta	ggctgaggac	cttgtggagt	2100
gtagctgac	cagctgatag	aggaactagc	caggtggggg	cctttccctt	tggatggggg	2160
gcataccga	cagttattct	ctccaagtgc	agacttacgg	acagcatata	attctcctgt	2220
caaggatgta	tgataaatatg	tacaaagtaa	ttccaactga	ggaagctcac	ctgatcccta	2280
gtgtccaggy	ttttactg	gggtctgtag	gacgagtag	gagtacttga	ataattgacc	2340
tgaagtccct	agacctgagg	ttccctagag	ttcaaacaga	tacagcatgg	tccagagtc	2400

<211> 557
 <212> DNA
 <213> Homo sapiens

<400> 384
 ggatcctcta gagcggccgc ctactactac taaattcgcg gccgcgtcga cgaagaagag 60
 aaagatgtgt tttgttttgg actctctgtg gtcccttcca atgctgtggg ttccaacca 120
 ggggaagggg cccttttgca ttgccaaagt ccataaccat gagcactact ctaccatggg 180
 tctgcctcct ggccaagcag gctggtttgc aagaatgaaa tgaatgattc tacagctagg 240
 acttaacctt gaaatggaaa gtcttgcaat cccatttgca ggatccgtct gtgcacatgc 300
 ctctgtagag agcagcattc ccagggacct tggaaacagt tggcactgta aggtgcttgc 360
 tccccaaagac acatcctaaa aggtgttgta atgggtgaaaa cgtcttcctt ctttattgcc 420
 ctttcttatt tatgtgaaca actgtttgtc tttttttgta tcttttttaa actgtaaagt 480
 tcaattgtga aaatgaatat catgcaaata aattatgcga tttttttttc aaagtaaaaa 540
 aaaaaaaaaa aaaaaaa 557

<210> 385
 <211> 337
 <212> DNA
 <213> Homo sapiens

<400> 385
 ttcccaggtg atgtgcgagg gaagacacat ttactatcct tgatggggct gattccttta 60
 gtttctctag cagcagatgg gttaggagga agtgaccaa gtggttgact cctatgtgca 120
 tctcaaagcc atctgctgtc ttcgagtacg gacacatcat cactcctgca ttgttgatca 180
 aaacgtggag gtgcttttcc tcagctaaga agcccttagc aaaagctcga atagacttag 240
 tatcagacag gtccagtttc cgcaccaaca cctgctggtt ccctgtcgtg gtctggatct 300
 ctttggecac caattcccc tttccacat cccggca 337

<210> 386
 <211> 300
 <212> DNA
 <213> Homo sapiens

<400> 386
 gggcccgcta ccggcccagg cccgcctcg cgagtcctcc tccccgggtg cctgcccgca 60
 gcccgctcgg ccagaggggt gggcgcgggg ctgcctctac cggctggcgg ctgtaactca 120
 ggcaccttgg ccgaaggct ctagcaagga ccaccgacc ccagccgcgg cggcggcggc 180
 ggggactttg cccggtgtgt ggggaggagc ggactgcgtg tccgcggacg ggcagcgaag 240
 atgttagcct tcgctgccag gaccgtggac cgatcccagg gctgtggtgt aacctcagcc 300

<210> 387
 <211> 537
 <212> DNA
 <213> Homo sapiens

<400> 387
 gggccgagtc gggcaccaag ggactctttg caggcttcct tctcggatc atcaaggctg 60
 ccccctcctg tgccatcatg atcagcacct atgagttcgg caaaagcttc ttccagaggc 120
 tgaaccagga ccggcttctg ggcggctgaa aggggcaagg aggcaaggac cccgtctctc 180
 ccacggatgg ggagagggca ggaggagacc cagccaagtg ctttttcctc agcactgagg 240
 gagggggctt gtttcccttc cctcccggcg acaagctcca gggcagggct gtccctctgg 300
 gcggcccagc acttcctcag acacaacttc ttctgctgc tccagtcgtg gggatcatca 360
 cttaccacac ccccaagttc aagaccaa atctccagctg cccccttcgt gtttccctgt 420

gtttgctgta gctgggcatg tctccaggaa ccaagaagcc ctcagcctgg tgtagtctcc 480
ctgacccttg ttaattcctt aagtctaaag atgatgaact tcaaaaaaaaa aaaaaaa 537

<210> 388

<211> 520

<212> DNA

<213> Homo sapiens

<400> 388

aggataatth ttaaaccaat caaatgaaaa aaacaaacaa acaaaaaagg aaatgtcatg 60
tgagggttaaa ccagtttgca ttcccctaata gtggaaaaag taaggaggact actcagcaact 120
gtttgaagat tgcctcttct acagcttctg agaatttgtt tatttcaactt gccaaagtga 180
ggacccccctc cccaacatgc ccagccccac ccctaagcat ggtcccttgt caccaggcaa 240
ccaggaaact gctacttggt gacctcacca gagaccagga gggtttggtt agctcacagg 300
acttccccca cccagaaga ttagcatccc atactagact cataactcaac tcaactaggc 360
tcatactcaa ttgatgggta ttagacaatt ccatttcttt ctgggttatta taaacagaaa 420
atctttcctc ttctcattac cagtaaaggc tcttggtatc tttctggttg aatgatttct 480
atgaacttgt cttattttta tgggtgggtt ttttctgtgt 520

<210> 389

<211> 365

<212> DNA

<213> Homo sapiens

<400> 389

cgttgccccg gtttgacaga aggaaaggcg gagcttattc aaagtctaga gggagtggag 60
gagtttaaggc tggatttcag atctgcctgg ttccagccgc agtgtgccct ctgctcccc 120
aacgactttc caaataatct caccagcgcc ttccagctca ggcgtcctag aagcgtcttg 180
aagcctatgg ccagctgtct ttgtgttccc tctcaccgc ctgtcctcac agctgagact 240
cccaggaaac cttcagacta ctttctctg cttcagcaa ggggcgttgc ccacattctc 300
tgagggtcag tggaagaacc tagactccca ttgctagagg tagaaagggg aagggtgctg 360
gggag 365

<210> 390

<211> 221

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(221)

<223> n = A,T,C or G

<400> 390

tgcctctcca tcttgcccc gacttctctg tcaggaaagt ggggatggac cccatctgca 60
tacacggntt ctcatgggtg tggaacatct ctgcttgccg tttcaggaag gcctctggct 120
gctctangag tctgancnga ntcgttgccc cantntgaca naaggaaagg cggagcttat 180
tcaaagtcta gagggagtgg aggagttaag gctggatttc a 221

<210> 391

<211> 325

<212> DNA

<213> Homo sapiens

tcttaacaac gaccgaaacc cattatTTac ataaacctcc attcggtaac catgttgaaa 300
gga 303

<210> 404
<211> 225
<212> DNA
<213> Homo sapiens

<400> 404
aagtgtaaCT tttaaaaatt tagtggattt tgaaaattct tagaggaaag taaaggaaaa 60
attgttaaat cactcattta cctttacatg gtgaaagtTc tctcttgatc ctacaaacag 120
acattttcca ctctgtgtttc catagtTgtt aagtgtatca gatgtgttg gcatgtgaat 180
ctccaagtgc ctgtgtaata aataaagtat ctttatTTca tTcat 225

<210> 405
<211> 334
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(334)
<223> n = A,T,C or G

<400> 405
gagctgttat actgtgagtt ctactaggaa atcatcaaT ctgaggggtg tctggaggac 60
ttcaatacac ctccccccat agtgaatcag cttccagggg gtccagtccc tctccttact 120
tcatccccat cccatgccaa aggaagaccT tccctccttg gctcacagcc ttctctaggc 180
ttcccagtgc ctccaggaca gagtgggtta tgttttcagc tccatccttg ctgtgagTgt 240
ctggtgcggt tgtgcctcca gcttctgctc agtgcTtcat ggacagtgtc cagcccatgt 300
cactctccac tctctcanng tggatccac cct 334

<210> 406
<211> 216
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(216)
<223> n = A,T,C or G

<400> 406
tttcatacct aatgagggag ttganatnac atnnaaccag gaaatgcatg gatctcaang 60
gaaacaaca cccaataaac tcggagtggc agactgacaa ctgtgagaca tgcacttgct 120
acnaaacaca aatttnatgt tgcacccttg tttctacacc tgtgggttat gacaaagaca 180
actgccaaag aatnttcaag aaggaggact gccant 216

<210> 407
<211> 413
<212> DNA
<213> Homo sapiens

<400> 407


```

cccagggacc ttggaaacag ttggcactgt aaggtgcttg ctccccaaga cacatcctaa 180
aaggtgttgt aatggtgaaa accgcttcct tctttattgc cccttcttat ttatgtgaac 240
nactggttgg ctttttttgn atctttttta aactggaaag ttcaattgng aaaatgaata 300
tcntgc                                           306

```

```

<210> 411
<211> 261
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(261)
<223> n = A,T,C or G

```

```

<400> 411
agagatattn cttaggtnaa agttcataga gttcccatga actatatgac tggccacaca 60
ggatcttttg tatttaagga ttctgagatt ttgcttgagc aggattagat aaggctgttc 120
tttaaagtgc tgaaatggaa cagatttcaa aaaaaaaccc cacaatctag ggtgggaaca 180
aggaaggaaa gatgtgaata ggctgatggg caaaaaacca atttaccat cagttccagc 240
cttctctcaa ggngaggcaa a                                           261

```

```

<210> 412
<211> 241
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(241)
<223> n = A,T,C or G

```

```

<400> 412
gttcaatggt acctgacatt tctacaacac cccactcacc gatgtattcg ttgccagtg 60
ggaacatacc agcctgaatt tggaaaaaat aattgtgttt cttgccagg aaatactacg 120
actgactttg atggctccac aaacataacc cagtgtaaaa acagaagatg tggaggggag 180
ctgggagatt tcaactgggta cattgaattc ccaaactacc cangcaatta cccagccaac 240
a                                           241

```

```

<210> 413
<211> 231
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(231)
<223> n = A,T,C or G

```

```

<400> 413
aactcttaca atccaagtga ctcatctgtg tgcttgaate ctttcactg tctcatctcc 60
ctcatccaag tttctagtag cttctctttg ttgtgaagga taatcaaact gaacaacaaa 120
aagtttactc tctctatttg gaacctaaaa actctcttct tcttgggtct gagggctcca 180
agaatccttg aatcanttct cagatcattg gggacaccan atcaggaacc t                                           231

```

<210> 414
 <211> 234
 <212> DNA
 <213> Homo sapiens

<400> 414
 actgtccatg aagcactgag cagaagctgg aggcacaacg caccagacac tcacagcaag 60
 gatggagctg aaaacataac ccactctgtc ctggaggcac tgggaagcct agagaaggct 120
 gtgagccaag gagggagggt cttcctttgg catgggatgg ggatgaagta aggagaggga 180
 ctggaccccc tggaagctga ttcactatgg ggggaggtgt attgaagtcc tcca 234

<210> 415
 <211> 217
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(217)
 <223> n = A,T,C or G

<400> 415
 gcataggatt aagactgagt atcttttcta cattctttta acttttctaag gggcacttct 60
 caaaacacag accaggtagc aaatctccac tgctctaagg ntctcaccac cacttttctca 120
 cacctagcaa tagtagaatt cagtctact tctgaggcca gaagaatggt tcagaaaaat 180
 antggattat aaaaaataac aattaagaaa aataatc 217

<210> 416
 <211> 213
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(213)
 <223> n = A,T,C or G

<400> 416
 atgcatatnt aaagganact gcctcgcttt tagaagacat ctggnctgct ctctgcatga 60
 ggcacagcag taaagctctt tgattcccag aatcaagaac tctccccttc agactattac 120
 cgaatgcaag gtgggtaatt gaaggccact aattgatgct caaatagaag gatattgact 180
 atattggaac agatggagtc tctactacaa aag 213

<210> 417
 <211> 303
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(303)
 <223> n = A,T,C or G

<210> 424
 <211> 370
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(370)
 <223> n = A,T,C or G

<400> 424
 gctcaaaaat ctttttactg ataggcatgg ctacacaatc attgactatt agaggccaga 60
 ggagaatgag gcttggcctg ggagccctgt gcctactaga agcacattag attatccatt 120
 cactgacaga acaggtcttt tttgggtcct tcttctccac cacgatatac ttgcagtcct 180
 ccttcttgaa gattctttgg cagttgtctt tgtcataacc cacaggtgta gaaacatcct 240
 ggttgaatct cctggaactc cctcattagg tatgaaatag catgatgcat tgcataaagt 300
 cacgaagggtg gcaaagatca caacgctgcc cagganaaca ttcattgtga taagcaggac 360
 tccgtcgacg 370

<210> 425
 <211> 216
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(216)
 <223> n = A,T,C or G

<400> 425
 aattgctatn ntattttttg ccactcaaaa taattaccaa aaaaaaaaaa tnttaaatga 60
 taacaacnca acatcaaggc aaananaaca ggaatggntg actntgcata aatnggccga 120
 anattatcca ttatnttaag ggttgacttc aggntacagc acacagacaa acatgcccag 180
 gaggntntca ggaccgctcg atgtnttntg aggagg 216

<210> 426
 <211> 596
 <212> DNA
 <213> Homo sapiens

<400> 426
 cttccagtga ggataaccct gttgccccgg gccgagggttc tccattaggc tctgattgat 60
 tggcagtcag tgatggaagg gtgttctgat cattccgact gccccaaggg tcgctggcca 120
 gctctctgtt ttgctgagtt ggcagtagga cctaatttgt taattaagag tagatggtga 180
 gctgtccttg tattttgatt aacctaattg ccttcccagc acgactcgga ttcagctgga 240
 gacatcacgg caacttttaa tgaaatgatt tgaagggcca ttaagaggca cttcccgtta 300
 ttaggcagtt catctgcact gataacttct tggcagctga gctggtcgga gctgtggccc 360
 aaacgcacac ttggcctttt gttttgagat acaactctta atcttttagt catgcttgag 420
 ggtggatggc cttttcagct ttaacceaat ttgactgcc ttggaagtgt agccaggaga 480
 atacactcat atactcgtgg gcttagaggc cacagcagat gtcattggtc tactgcctga 540
 gtcccgcgtgg tcccatccca ggaccttcca tcggcgagta cctgggagcc cgtgct 596

<210> 427
 <211> 107

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(107)
<223> n = A,T,C or G

<400> 427
gaagaattca agttaggttt attcaaaggg cttacngaga atcctanacc caggncccag 60
cccgaggagca gccttanaga gtcctgttt gactgcccgg ctcagng 107

<210> 428
<211> 38
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(38)
<223> n = A,T,C or G

<400> 428
gaacttcna anaangactt tattcactat tttacatt 38

<210> 429
<211> 544
<212> DNA
<213> Homo sapiens

<400> 429
ctttgctgga cggaataaaa gtggacgcaa gcatgacctc ctgatgaggg cgctgcattt 60
attgaagagc ggctgcagcc ctgcggttca gattaaaatc cgagaattgt atagacgccg 120
atatccacga actcttgaag gactttctga tttatccaca atcaaatcat cgggttttcag 180
tttggatggt ggctcatcac ctgtagaacc tgacttggcc gtggctggaa tccactcgtt 240
gccttcact tcagttacac ctactcacc atcctctcct gttggttctg tgetgettca 300
agatactaag cccacatttg agatgcagca gccatctccc ccaattcctc ctgtccatcc 360
tgatgtgcag ttaaaaaatc tgccctttta tgatgtcctt gatgttctca tcaagcccac 420
gagtttagtt caaagcagta ttcagcgatt tcaagagaag ttttttattt ttgctttgac 480
acctcaacaa gttagagaga tatgcatatc cagggatttt ttgccagggtg gtaggagaga 540
ttat 544

<210> 430
<211> 507
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(507)
<223> n = A,T,C or G

<400> 430
cttatcncaa tggggctccc aaacttggtt gtgcagtgga aactccgggg gaattttgaa 60

```

gaacactgac acccatcttc caccocgaca ctctgattta attgggctgc agtgagaaca 120
gagcatcaat ttaaaaaagct gcccagaatg ttntcctggg cagcgttgtg atctttgccn 180
ccttcgtgac tttatgcaat gcatcatgct atttcatacc taatgaggga gttccaggag 240
attcaaccag gatgtttcta cncctgtggg ttatgacaaa gacaactgcc aaagaatntt 300
caagaaggag gactgcaagt atatcgtggt ggagaagaag gacccaaaaa agacctgttc 360
tgtcagtga tggataatct aatgtgcttc tagtaggcac agggctccca ggccaggcct 420
cattctctc tggcctctaa tagtcaatga ttgtgtagcc atgcctatca gtaaaaagat 480
ttttgagcaa aaaaaaaaaa aaaaaaa 507

```

```

<210> 431
<211> 392
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(392)
<223> n = A,T,C or G

```

```

<400> 431
gaaaattcag aatggataaa aacaaatgaa gtacaaaata tttcagattt acatagcgat 60
aaacaagaaa gcacttatca ggaggactta caaatggaag tacactctan aaccatcatc 120
tatcatggct aaatgtgaga ttagcacagc tgtattattt gtacattgca aacacctaga 180
aagagatggg aaacaaaatc ccaggagttt tgtgtgtgga gtccctgggtt ttccaacaga 240
catcattcca gcattctgag attaggngga ttggggatca ttctggagtt ggaatgttca 300
acaaaagtga tgttgttagg taaaatgtac aacttctgga tctatgcaga cattgaaggt 360
gcaatgagtc tggcttttac tctgctgttt ct 392

```

```

<210> 432
<211> 387
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(387)
<223> n = A,T,C or G

```

```

<400> 432
ggtatccnta cataatcaaa tatagctgta gtacatgttt tcattggngt agattaccac 60
aaatgcaagg caacatgtgt agatctcttg tcttattctt ttgtctataa tactgtattg 120
ngtagtccaa gctctcgga gtccagccac tgngaaacat gctcccttta gattaacctc 180
gtggacnctn ttgttgnatt gtctgaactg tagngccctg tattttgctt ctgtctgnga 240
attctgttgc ttctggggca ttctcttng atgcagagga ccaccacaca gatgacagca 300
atctgaattg ntccaatcac agctgcgatt aagacatact gaaatcgtac aggaccggga 360
acaacgtata gaacactgga gtccttt 387

```

```

<210> 433
<211> 281
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature

```

<222> (1)...(281)

<223> n = A,T,C or G

<400> 433

```
ttcaactagc anagaanact gcttcagggg gtgtaaaatg aaaggcttcc acgcagttat 60
ctgattaaag aacactaaga gagggacaag gctagaagcc gcaggatgtc tacactatag 120
caggcnctat ttgggttggc tggaggagct gtggaaaaca tggagagatt ggcgctggag 180
atcgccgtgg ctattcctcn ttgntattac accagngagg ntctctgtnt gccactgggt 240
tnnaaaaccg ntatacaata atgatagaat aggacacaca t 281
```

<210> 434

<211> 484

<212> DNA

<213> Homo sapiens

<400> 434

```
ttttaaaata agcatttagt gctcagtcct tactgagtac tctttctctc cctcctctctg 60
aatttaattc tttcaacttg caatttgcaa ggattacaca tttcactgtg atgtatattg 120
tggtgcaaaa aaaaaaaagt gtctttgttt aaaattactt gggttggtgaa tccatcttgc 180
tttttcccca ttggaactag tcattaaccc atctctgaac tggtagaaaa acatctgaag 240
agctagtcta tcagcatctg acaggtgaat tggatggttc tcagaaccat ttcaccaga 300
cagcctgttt ctatcctgtt taataaatta gtttgggttc tctacatgca taacaaaccc 360
tgctccaatc tgtcacataa aagtctgtga cttgaagttt agtcagcacc cccaccaaac 420
tttatttttc tatgtgtttt ttgcaacata tgagtgtttt gaaaataaag taccatgtc 480
ttta 484
```

<210> 435

<211> 424

<212> DNA

<213> Homo sapiens

<400> 435

```
gogccgctca gagcaggtca ctttctgcct tccacgtcct ccttcaagga agcccatgt 60
gggtagcttt caatatcgca ggttcttact cctctgcctc tataagctca aaccaccaa 120
cgatcgggca agtaaacccc ctccctcgcc gacttcggaa ctggcgagag ttcagcgcag 180
atgggcctgt ggggaggggg caagatagat gagggggagc ggcatgggtc ggggtgaacc 240
cttgagagaga ggaaaaaggc cacaagaggg gctgccaccg ccactaacgg agatggccct 300
ggtagagacc tttgggggtc tggaacctct ggactcccca tgctctaact cccacactct 360
gctatcagaa acttaaactt gaggattttc tctgtttttc actcgcaata aattcagagc 420
aaac 424
```

<210> 436

<211> 667

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(667)

<223> n = A,T,C or G

<400> 436

```
accttgggaa nactctcaca atataaaggg tcgtagactt tactccaaat tccaaaaagg 60
tcctggccat gtaatcctga aagttttccc aaggtagcta taaaatcctt ataagggtgc 120
```



```

agcctcttct ggaattcctc tgatttcaaa gtctcactct caagttcttg aaaacgaggg 180
cagttcctga aaggcaggta tagcaactga tcttcagaaa gaggaactgt gtgcaccggg 240
atgggctgcc agagtaggat aggattccag atgctgacac cttctggggg aaacagggct 300
gccaggtttg tcatagcact catcaaagtc cggtcacagt ctgtgcttcg aatataaacc 360
tgttcatgtt tataggactc attcaagaat tttctatata tctttcttat atactctcca 420
agttcataat gctgctccat gccagctgg gtgagttggc caaatccttg tggccatgag 480
gattccttta tggggtcagt gggaaagggt tcaatgggac ttcggtctcc atgccgaaac 540
accaaagtca caaacttcaa ctcttggt agtacacttc ggtctagcca gaaaaaagc 600
agaaacaaga agccaaggct aaggcttgct gccctgccag gaggaggggt gcagctctca 660
tgttgag                                           667

```

<210> 437

<211> 693

<212> DNA

<213> Homo sapiens

<400> 437

```

ctacgtctca accctcattt ttaggtaagg aatcttaagt ccaaagatat taagtgactc 60
acacagccag gtaaggaaaag ctggattggc acactaggac tctaccatac cgggttttgt 120
taaagctcag gttaggaggc tgataagctt ggaaggaaact tcagacagct ttttcagatc 180
ataaaagata attcttagcc catgttcttc tccagagcag acctgaaatg acagcacagc 240
aggtactcct ctattttcac cctcttggt tctactctct ggcagtcaga cctgtgggag 300
gccatgggag aaagcagctc tctggatgtt tgtacagatc atggactatt ctctgtggac 360
catttctcca ggttacccta ggtgtcacta ttgggggggac agccagcacc tttagctttc 420
atgtgagttt ctgtctgtct tcagttagagg aaacttttgc tcttcacact tcacatctga 480
acacctaact gctgttgctc ctgaggtggt gaaagacaga tatagagctt acagtattta 540
tcctattttc aggcactgag ggctgtgggg taccttgtgg tgccaaaaca gatcctgttt 600
taaggacatg ttgcttcaga gatgtctgta actatctggg ggcctctgtt gctctttacc 660
ctgcatcatg tgctctcttg gctgaaaatg acc                                           693

```

<210> 438

<211> 360

<212> DNA

<213> Homo sapiens

<400> 438

```

ctgcttatca caatgaatgt tctcctgggc agcgttggtga tctttgccac cttcgtgact 60
ttatgcaatg catcatgcta tttcatacct aatgagggag ttccaggaga ttcaaccagg 120
atgtttctac acctgtgggt tatgacaaag acaactgcca aagaatcttc aagaaggagg 180
actgcaagta tatctgggtg agaagaagga ccaaaaaaag acctgttctg tcagtgaatg 240
gataatctaa tgtgcttcta gtaggcacag ggctcccagg ccaggcctca ttctcctctg 300
gcctctaata gtcaataatt gtgtagccat gcctatcagt aaaaagattt ttgagcaaac 360

```

<210> 439

<211> 431

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (431)

<223> n = A,T,C or G

<400> 439

```
<210> 440
<211> 523
<212> DNA
<213> Homo sapiens
```

```
<210> 441
<211> 430
<212> DNA
<213> Homo sapiens
```

```
<210> 442
<211> 362
<212> DNA
<213> Homo sapiens
```

```

<400> 442
ctaaggaatt agtagtggtc ccatcacttg tttggagtg tctattctaa aagattttga 60
tttctctggaa tgacaattat attttaactt tgggtggggga aagagttata ggaccacagt 120
cttcacttct gatacttgta aattaatctt ttattgcact tgttttgacc attaatgctat 180
atgttttagaa atggtcattt tacggaaaaa ttagaaaaat tctgataata gtgcagaata 240
aatgaattaa tgttttactt aattttatatt gaactgtcaa tgacaaataa aaattctttt 300
tgattatttt ttgttttcat ttaccagaat aaaaactaag aattaaaagt ttgattacag 360
tc
362

```

<400> 445

```

catgtttatg nttttggatt actttgggca cctagtgttt ctaaatcgtc tatcattctt 60
ttctgttttt caaaagcaga gatggccaga gtctcaacaa actgtatctt caagtctttg 120
tgaaattctt tgcattgtggc agattattgg atgtagtctt ctttaactag catataaatc 180
tggtgtgttt cagataaatg aacagcaaaa tgtggtggaa ttaccatttg gaacattgtg 240
aatgaaaaat tgtgtctcta gattatgtaa caaataacta tttcctaacc attgatcttt 300
ggatttttat aatcctactc acaaatgact aggcctctcc tcttgatatt tgaagcagtg 360
tggtgtctgg attgataaaa aaaaaaaaaa tgcgcgcggc cgcgaattta gtag 414

```

```

<210> 446
<211> 631
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(631)
<223> n = A,T,C or G

```

```

<400> 446
acaaattaga anaaagtgcc agagaacacc acataccttg tccggaacat tacaatggct 60
tctgcatgca tgggaagtgt gagcattcta tcaatatgca ggagccatct tgcagggtgtg 120
atgctgggta tactggacaa cactgtgaaa aaaaggacta cagtgttcta tacgttggtc 180
ccggctctgt acgatttcag tatgtcttaa tgcagctgtg gattggaaca attcagattg 240
ctgtcatctg tgtggtgggc ctctgcatca caagggccaa actttaggta atagcattgg 300
actgagattt gtaaactttc caaccttcca ggaaatgccc cagaagcaac agaattcaca 360
gacagaagca aaatacaggg cactacagtt cagacaatac aacaagagcg tccacgaggt 420
taatctaaag ggagcatgtt tcacagtggc tggactaccg agagcttgga ctacacaata 480
cagtattata gacaaaagaa taagacaaga gatctacaca tgttgccctg catttggtgtg 540
aatctacacc aatgaaaaca tgtactacag ctatatattga ttatgtatgg atatatattga 600
aatagtatac attgtcttga tgttttttct g 631

```

```

<210> 447
<211> 585
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(585)
<223> n = A,T,C or G

```

```

<400> 447
ccttgggaaa antntcacia tataaagggt cgtagacttt actccaaatt ccaaaaagggt 60
cctggccatg taatcctgaa agttttccca aggtagctat aaaatcetta taagggtgca 120
gcctcttctg gaattcctct gatttcaaag tctcactctc aagtcttga aaacgagggc 180
agttcctgaa aggcaggat agcaactgat cttcagaaa aggaactgtg tgcaccggga 240
tggtgtgcca gagtaggata ggattccaga tgcagacacc ttctggggga aacagggctg 300
ccagggttgt catagcactc atcaaagtcc ggtcaacgtc tgtgcttga atataaacct 360
gttcatgttt ataggactca ttcaagaatt ttctatatct ctttcttata tactctccaa 420
gttcataatg ctgctccatg cccagctggg tgagttggcc aaatccttgt ggccatgagg 480
attcctttat ggggtcagtg ggaaagggtg caatgggact tccgtctcca tgccgaaaca 540
ccaaagtcac aaacttcaac tccttggcta gtacacttcg gtcta 585

```

```

<210> 448

```


<212> DNA

<400> 454

<210> 455

<212> DNA

<213> Homo sapiens

taccaaagag	ggcataataa	tcagtctcac	agtaggggtc	accatcctcc	aagtgaaaaa	60
cattgttccg	aatgggcttt	ccacaggcta	cacacacaaa	acaggaaaca	tgccaagttt	120
gtttcaacgc	attgatgact	tctccaagga	tcttcctttg	gcctcgacca	cattcagggg	180
caaagaattt	ctcatagcac	agctcacaat	acaggggctcc	tttctcctct	a	231

<211> 231

<212> DNA

<213> Homo sapiens

ttggcaggta	cccttacaaa	gaagacacca	taccttatgc	gttattaggt	ggaataatca	60
ttccattcag	tattatcggt	attattcttg	gagaaacct	gtctgtttac	tgtaaccttt	120
tgcactcaaa	ttcctttatc	aggaataact	acatagccac	tatttacaaa	gccattggaa	180
cctttttatt	tggtgcagct	gctagtcagt	ccctgactga	cattgccaaag	t	231

<210> 457

<211> 231

<212> DNA

<213> Homo sapiens

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (231)$

<223> n = A, T, C or G

<400> 457

cgaggtagcc	aggggtctga	aaatctctnn	tttantagtc	gatagcaaaa	ttgttcac	60
gcattcccta	atatgatctt	gctataatta	gattttttctc	cattagagtt	catacagttt	120
tatttgattt	tattagcaat	ctcttttcaga	agacccttga	gatcattaag	ctttgtatcc	180
agttgtctaa	atcgatgcct	catttcctct	gagggtgtcgc	tggcttttgt	g	231

<210> 458

<211> 231

<212> DNA

<213> Homo sapiens

<400> 458

aggtctgggt cccccactt ccactccct ctactctctc taggactggg ctgggccaaag 60
agaagagggg tggtaggga agccgttgag acctgaagcc ccaccteta ccttccttca 120

```
<210> 459
<211> 231
<212> DNA
<213> Homo sapiens
```

```
<210> 460
<211> 231
<212> DNA
<213> Homo sapiens
```

```
<210> 461
<211> 231
<212> DNA
<213> Homo sapiens
```

```
<210> 462
<211> 231
<212> DNA
<213> Homo sapiens
```

```
<210> 463
<211> 231
<212> DNA
<213> Homo sapiens
```

<400> 463
tactccagcc tgggtgacaga gcgagaccct atcaccgccc cccaccccac caaaaaaaaa 60

<213> Homo sapiens

<400> 468

```

cattgtgttg ggagaaaaac agaggggaga tttgtgtggc tgcagccgag ggagaccagg 60
aagatctgca tgggtgggaag gacctgatga tacagagttt gataggagac aattaaaggc 120
tggaaggcac tggatgcctg atgatgaagt ggactttcaa actggggcac tactgaaacg 180
atgggatggc cagagacaca ggagatgagt tggagcaagc tcaataacaa agtgggttcaa 240
cgaggacttg gaattgcatg gagctggagc tgaagtttag cccaattggt tactagttag 300
gtgaatgtgg atgattggat gatcatttct catctctgag cctcagggtc cccatccata 360
aaatgggata cacagtatga tctataaagt gggatatagt atgatctact tcaactgggt 420
atgtgaagga tgaattgaga taatttattt cagggtgcct gaacaatgcc cagattagta 480
catttggtgg aactgagaaa tggcataaca ccaaatttaa tatatgtcag atgttactat 540
gattatcatt caatctcata gttttgtcat ggcccaattt atcctcactt gtgcctcaac 600
aaattgaact gttaacaaag gaatctctgg tcctgggtaa tggctgagca ccactgagca 660
tttccattcc agttggcttc ttgggtttgc tagctgcac actagtcac tttaataaat 720
gaagttttaa catttctcca gtgatttttt tatctcacct ttgaagatac tatgttatgt 780
gattaaataa agaacttgag aagaacaggc ttcattaaac ataaaatcaa tgtagacgca 840
aattttctgg atgggcaata cttatgttca caggaaatgc tttaaaatat gcagaagata 900
attaaatggc aatggacaaa gtgaaaaact tagacttttt tttttttttt ggaagtatct 960
ggatgttcct tagtcactta aaggagaact gaaaaatagc agtgagtacc acataatcca 1020
acctgtgaga ttaaggctct ttgtggggaa ggacaaaagat ctgtaaattt acagtttccct 1080
tccaaagcca acgtcgaatt ttgaaacata tcaaagctct tcttcaagac aaataatcta 1140
tagtacatct ttcttatggg atgcacttat gaaaaatggt ggctgtcaac atctagtcac 1200
tttagctctc aaaatggttc attttaagag aaagtttttag aatctcatat ttattcctgt 1260
ggaaggacag cattgtggct tggactttat aaggtcttta ttcaactaaa taggtgagaa 1320
ataagaaagg ctgctgactt taccatctga ggccacacat ctgctgaaat ggagataatt 1380
aacatcacta gaaacagcaa gatgacaata taatgtctaa gtagtgacat gtttttgcac 1440
atttccagcc cctttaaata tccacacaca caggaagcac aaaaggaagc acagagatcc 1500
ctgggagaaa tgcccggccg ccatcttggg tcacgatga gcctcgccct gtgcctgggtc 1560
ccgcttgtga ggggaaggaca ttagaaaatg aattgatgtg ttccttaaag gatgggcagg 1620
aaaacagatc ctgttgtgga tatttatttg aacgggatta cagatttgaa atgaagtcac 1680
aaagtgagca ttaccaatga gaggaaaaca gacgagaaaa tcttgatggc ttcacaagac 1740
atgcaacaaa caaaatggaa tactgtgat acatgaggca gccaaagctgg ggaggagata 1800
accacggggc agagggtcag gattctggcc ctgctgccta aactgtgcgt tcataaccaa 1860
atcatttcat atttctaacc ctcaaaacaa agctgttgta atatctgac tctacgggtc 1920
cttctgggcc caacattctc catatatcca gccacactca tttttaatat ttagttccca 1980
gatctgtact gtgacctttc tacactgtag aataacatta ctcattttgt tcaaagacc 2040
ttcgtgttgc tgccataat gttagctgact gtttttccca aggagtgttc tggcccagg 2100
gatctgtgaa caggctggga agcatctcaa gatctttcca gggttatact tactagcaca 2160
cagcatgac attacggagt gaattatcta atcaacatca tcctcagtgt ctttgcccat 2220
actgaaattc atttccact tttgtgcccc ttctcaagac ctcaaaatgt cattccatta 2280
atatcacagg attaactttt ttttttaacc tggaagaatt caatgttaca tgcagctatg 2340
ggaatttaaa tacatatatt gttttccagt gcaaagatga ctaagtcctt tatccctccc 2400
ctttgtttga ttttttttcc agtataaagt taaaatgctt agccttgtag tgaggctgta 2460
tacagccaca gcctctcccc atccctccag ccttatctgt catcaccatc aaccctccc 2520
atgcacctaa acaaaatcta acttgtaatt ccttgaacat gtcaggcata cattattcct 2580
tctgcctgag aagctcttcc ttgtctctta aatctagaat gatgtaaaat tttgaataag 2640
ttgactatct tacttcatgc aaagaaggga cacatatgag attcatcatc acatgagaca 2700
gcaataacta aaagtgtaat ttgattataa gagtttagat aaatatatga aatgcaagag 2760
ccacagaggg aatgtttatg gggcacgttt gtaagcctgg gatgtgaagc aaaggcagg 2820
aacctcatag tatcttataa aatatacttc atttctctat ctctatcaca atatccaaca 2880
agcttttccac agaattcatg cagtgc aaat aacctttatc catttctagg 2940
tgagtgcgct ttagaatttt ggcaaatcat actggctcact tatctcaact ttgagatgtg 3000
tttgtccttg tagttaattg aaagaaatag ggcactcttg tgagccactt tagggttcac 3060
tcctgggcaat aaagaattta caaagagcaa aaaaaaaaaa aaaaaaaaaa aa 3112

```

<210> 469
 <211> 2229
 <212> DNA
 <213> Homo sapiens

<400> 469
 agctctttgt aaattcttta ttgccaggag tgaaccctaa agtgggtcac aagagtgtcc 60
 tatttctttc aattaactac aaggacaaac acatctcaaa gttgagataa gtgaccagta 120
 tgatttgcca aaattctaaa gcgcactcac catgaaatgg ataaagggtta cctttgggga 180
 tttgcactgc atgaattctg tgaaaagctt gttggatatt gtgatagaga tagagaaatg 240
 aagtatatta tataagatac tatgagggtc cctgcctttg cttcacatcc caggcttaca 300
 aacgtgtccc ataaacattc cctctgtggc tcttgcaatt catatattta tctaaactct 360
 tataatcaaa tacactttta gtatttgctg tctcatgtga tgatgaatct catatgtgtc 420
 ccttctttgc atgaagtaag atagtcaact tattcaaaac ttacatcat tctagattta 480
 agagacaagg aagagcttct caggcagaag gaataatgta tgctgacat gttcaaggaa 540
 ttacaagtta gattttgttt aggtgcatgg gaggggttga tgggtgatgac agataaggct 600
 ggagggatgg ggagaggctg tggctgtata cagcctcagt acaaggctaa gcattttaac 660
 tttatactgg aaaaaaaatc aaacaaaggg gagggataaa ggacttagtc atctttgcac 720
 tggaaaacaa aatatgtaat taaattccca tagctgcatg taacattgaa ttcttcagg 780
 ttaaaaaaaa agttaatcct gtgatattaa tggaaatgaca ttttgaggtc ttgagaatgg 840
 gcacaaaagt gggaaatgaa tttcagtatg ggcaaaagaca ctgaggatga tgttgattag 900
 ataattcact ccgtaatgat catgctgtgt gctagtaagt ataaccctgg aaagatcttg 960
 agatgcttcc cagcctgttc acagatcccc tgggccagaa cactccttag gaaaaacagt 1020
 cagctacata ttaggcagca acacgaaggg tctttgaaca aaatgagtaa tgttattcta 1080
 cagtgtagaa aggtcacagt acagatctgg gaactaaata ttaaaaatga gtgtggctgg 1140
 atatatggag aatgttgggc ccagaaggaa ccgtagagat cagatattac aacagctttg 1200
 ttttgagggt tagaaatatg aaatgatttg gttatgaacg cacagtttag gcagcagggc 1260
 cagaatcctg accctctgcc ccgtgggtat ctctcccca gcttggctgc ctcatgtcat 1320
 cacagtattc cattttgttt gttgcatgtc ttgtgaagcc atcaagattt tctcgtctgt 1380
 tttcctctca ttggtaatgc tcactttgtg acttcatttc aaatctgtaa tcccgttcaa 1440
 ataaatatcc acaacaggat ctgttttctt gccatcctt taagggaacac atcaattcat 1500
 tttctaattg ctttccctca caagcgggac caggcacagg gcgaggctca togatgacc 1560
 aagatggcgg ccgggcattt ctcccaggga tctctgtgct tctttttgtg ctctctgtgt 1620
 gtgtggatat ttaaaggggc tggaaatgtg caaaaacatg tcactactta gacattatat 1680
 tgtcatcttg ctgtttctag tgatgttaat tatctccatt tcagcagatg tgtggcctca 1740
 gatggtaaa gtcagcagct ttcttatttc tcacttgga atacatacga ccatttgagg 1800
 agacaaatgg caaggtgtca gcataccctg aacttgagtt gagagctaca cacaatatta 1860
 ttggtttccg agcatcaca acacctctc tgtttcttca ctgggcacag aattttaata 1920
 cttatttcag tgggctgttg gcaggaacaa atgaagcaat ctacataaag tcaactagtgc 1980
 agtgctgac acacaccatt ctcttgaggt cccctctaga gatccacag gtcatatgac 2040
 ttcttgggga gcagtggctc acacctgtaa tcccagcact ttgggaggct gaggcagggtg 2100
 ggtcacctga ggtcaggagt tcaagaccag cctggccaat atggtgaaac cccatctcta 2160
 ctaaaaatac aaaaattagc tgggcgtgct ggtgcatgcc tgtaatccca gcccacacac 2220
 aatggaatt 2229

<210> 470
 <211> 2426
 <212> DNA
 <213> Homo sapiens

<400> 470
 gtaaatctt tattgccagg agtgaacctt aaagtggctc acaagagtgc cctatttctt 60
 tcaattaact acaaggacaa acacatctca aagttgagat aagtgaccag tatgatttgc 120

caaaattcta aagcgcactc accatgaaat ggataaaggt tacctttggg gatttgcact 180
 gcatgaattc tgtgaaaagc ttgttggata ttgtgataga gatagagaaa tgaagtatat 240
 tatataagat actatgaggt tccctgcctt tgcttcacat cccaggctta caaacgtgcc 300
 ccataaacat tccctctgtg gctcttgcac ttcataatatt tatctaaact cttataatca 360
 aattacactt ttagtatttg ctgtctcatg tgatgatgaa tctcatatgt gtcccttctt 420
 tgcataagat aagatagtc aacttattcaa aactttacat cattctagat ttaagagaca 480
 aggaagagct tctcaggcag aaggaataat gtatgcctga catgttcaag gaattacaag 540
 ttagattttg tttagggtgca tgggaggggt tgatgggtgat gacagataag gctggaggga 600
 tggggagagg ctgtggctgt atacagcctc agtacaaggc taagcatttt aactttatac 660
 tggaaaaaaa atcaaacaaa ggggagggat aaaggactta gtcacttttg cactggaaaa 720
 caaaatatgt aattaaattc ccatagctgc atgtaacatt gaattcttcc aggttaaaaa 780
 aaaaagttaa tccgtgtgata ttaatggaat gacattttga ggtcttgaga atgggcacaa 840
 aagtgggaaa tgaatttcag tatgggcaaa gacactgagg atgatgttga ttagataatt 900
 cactccgtaa tgatcatgct gtgtgctagt aagtataacc ctggaaagat cttgagatgc 960
 ttcccagcct gttcacagat cccctgggct agaactctc ttaggaaaaa cagtcagcta 1020
 catattaggc agcaacacga aggtctttg aacaaaaatga gtaatgttat tctacagtgt 1080
 agaaaggtca cagtacagat ctgggaacta aatattaaaa atgagtgttg ctggatatat 1140
 ggagaatggt gggcccagaa ggaaccgtag agatcagata ttacaacagc tttgttttga 1200
 gggtagaaaa tatgaaatga tttggttatg aacgcacagt ttaggcagca gggccagaat 1260
 cctgaccctc tgcccctgtg tttctctctc cccagcttgg ctgcctcatg tcatcacagt 1320
 attccatttt gtttgttgca tgtcttgtga agccatcaag attttctcgt ctgttttctt 1380
 ctcatgtgta atgtcactt tgtgacttca tttcaaatct gtaatcccg tcaataaat 1440
 atccacaaca ggatctgttt tctgcccct cctttaagga acacatcaat tcattttcta 1500
 atgtccttcc ctcaacaagc ggaccaggca cagggcgagg ctcatcgatg acccaagatg 1560
 gcggccgggc atttctccca gggatctctg tgcttcttct tgtgcttctt gtgtgtgtgg 1620
 atattttaaag gggctggaat tgtgcaaaaa catgtcacta cttagacatt atattgtcat 1680
 cttgctgttt ctagtgtatg taattatctc catttcagca gatgtgtggc ctcatatggt 1740
 aaagtcagca gcctttctta tttctcacct ggaataacat acgaccattt gaggagacaa 1800
 atggcaagggt gtcagcatac cctgaacttg agttgagagc tacacacaat attattggtt 1860
 tccgagcacc acaaacaccc tctctgtttc ttcactgggc acagaatttt aatacttatt 1920
 tcagtgggct gttggcagga acaaatgaag caatctacat aaagtcacta gtgcagtgc 1980
 tgacacacac cattctcttg aggtcccctc tagagatccc acaggtcata tgacttcttg 2040
 gggagcagtg gctcacacct gtaatcccag cactttggga ggctgaggca ggtgggtcac 2100
 ctgagggtcag gatttcaaga ccagcctggc caatatggtg aaaccccatc tctactaaaa 2160
 atacaaaaat tagctgggct tgctgggtga tgctgtaat cccagctact tgggaggctg 2220
 aggagggaga attgtgtgaa catgggaggc ggaggttgca gtgagctgta attgtgccat 2280
 tgcaactcgaa cctgggagac agagtgggac tctgtttcca aaaaacaaac aaacaaaaaa 2340
 ggcatagtca gatacaacgt gggtagggat tgtaaataga agcaggatat aaagggcatg 2400
 gggtagcggg tttgcccac acaatg 2426

<210> 471

<211> 812

<212> DNA

<213> Homo sapiens

<400> 471

gaacaaaatg agtaatgtta ttctacagt tagaaaaggtc acagtacaga tctgggaact 60
 aaatattaaa aatgagtgtg gctggatata tggagaatgt tgggcccaga aggaaccgta 120
 gagatcagat attacaacag ctttgttttg agggtagaa atatgaaatg atttggttat 180
 gaacgcacag tttaggcagc agggccagaa tccctgaccct ctgcccctgt gttatctcct 240
 cccagcttg gctgctcat gtcacacag tattccattt tgtttgttg atgtcttgtg 300
 aagccatcaa gattttctcg tctgttttcc tctcatttgt aatgctcact ttgtgacttc 360
 atttcaaata tgtaatcccg ttcaaataaa tatccacaac aggatctgtt ttctgccc 420
 tcttttaagg aacacatcaa ttcattttct aatgtccttc cctcacagc gggaccaggc 480

```
<210> 472
<211> 515
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(515)
<223> n = A,T,C or G
```

```
<210> 473
<211> 5829
<212> DNA
<213> Homo sapiens
```

<400> 473						
cgcatgccgg	ggaagcccaa	gctggctcga	agagccacca	gccacctgtg	caagggtggg	60
cctggaccag	ttggaccagc	caccaagctc	acctactcaa	ggaagcaggg	atggccaggt	120
tgcaacagcc	tgagtggctg	ccacctgata	gctgatggag	cagaggcctg	aggaaaatca	180
gatggcacat	ttagctcttt	aatggatctt	aagttaattt	ttctataaag	cacatggcac	240
cagtccatgc	ctcagagctc	gtatggcact	gcggaaccaca	gcaggccgag	ttcccaggat	300
tgccatccag	gggggccttc	tgtagccctg	gccagacctt	gcagaggttg	ctgggtgctc	360
tttgagcgag	ctcggcctcc	ctggcatgca	caggccccag	gtactgacac	gctgctctga	420
gtgagcttgt	cctgccttgg	ctgccaccta	actgctgatg	gagcagcggc	cttaggaaaa	480
gcaaagtggc	ctgtagccca	actttagggt	agaagaagat	gtaccatgtc	cggccgctag	540
ttggtgactg	gtgcacctgc	tcctggcgta	cccttgccaga	ggtgggtggt	tgctcttttg	600
ccagcttggc	cttgccctggc	atgcacaagc	ctcagtgcaa	caactgtcct	acaaatggag	660
acacagagag	gaaacaagca	gcgggctcag	gagcaggggtg	tgtgctgcct	ttggggctcc	720
agtccatgcc	tcgggtcgta	tggtactgca	ggcttctttg	ttgccaaagag	gcggaccaca	780
ggccttcttg	aggaggactt	tacgttcaag	tcgagaaaagc	agccaaaatt	accatccatg	840
agactaagcc	ttctgtggcc	ctggcagagc	ttaaaatttg	tgccaaggca	ggacaagctc	900
actcggagca	gcgtgtcagt	agctggggcc	tcgtcatgcc	gggcagggcc	gggctggctg	960
aaggagcaac	cagccacctc	tgcaaggggtg	cgcctagtgc	aggcggagca	tcaccacctt	1020
caccgctcgc	aggaagtggg	gatggccagg	ttcccacagc	ctgagtgtct	gccaccttat	1080
tgctgatgga	gcagaggcct	taagaaaagc	agatggcact	gtggccctac	ctttagggtg	1140
gaagaagtga	tgtacatgtc	cggacgctaa	ttggtgactg	gtacaccggc	tcctgtctaca	1200


```

gccgacaggg tgaatgcctt cttgcttgtc ttctggggaa tcagagagag tccgttttgt 1440
ggtagggaaga gtgtggctgt gtactttgaa ctctgtataa ttctctgact catgtccaca 1500
aaaccaacag ttttgtgaat gtgtctggag gcaagggag ggccactcag gatctatgtt 1560
gaaggggaaga ggcctggggc tggagtattc gctt 1594

```

<210> 475

<211> 2414

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33)

<223> n=A,T,C or G

<400> 475

```

cccaacacaa tggctttata agaatgcttc acntgtgaaa aacaaatata aaagtcttct 60
tgtagattat ttttaaggac aaatctttat tccatgttta atttatttag ctttccctgt 120
agctaataat tcatgctgaa cacattttaa atgctgtaaa tgtagataat gtaatttatg 180
tatcattaat gcctcttttag tagtttagag aaaacgtcaa aagaaatggc cccagaataa 240
gcttcttgat ttgtaaaatt ctatgtcatt ggctcaaatt tgtatagtat ctcaaaatat 300
aaatatatag acatctcaga taatatattt gaaatagcaa attcctgtta gaaaataata 360
gtacttaact agatgagaat aacaggctcg cattatttga attgtctcct attcgttttt 420
catttgttgt gttactcatg ttttacttat ggggggatat atataacttc cgctgttttc 480
agaagtattg tatgcagtca gtatgagaat gcaatttaag ttctcttgat gctttttcac 540
acttctatta ctagaaataa gaatacagta atattggcaa agaaaattga ccagttcaat 600
aaaatttttt agtaaactctg attgaaaata aacattgctt atggctttct tacatcaata 660
ttgttatgtc ctagacacct tatctgaaat tacggcttca aaattctaata tatgtgcaaa 720
tgtgtaaaaa atcaataactt tatgttcaag ctggggcctc ttcaggcgctc ctgggctgag 780
agagaaagat gctagctccg caagccgggg agggaaacacc gccacattgt tacatggaca 840
caccgccacg tggacacatg accagactca catgtacaga cacacggaga cattaccaca 900
tggagacacc gtcacacagt cacacgagca cactggcata gtcacatgga cggacacaca 960
gacatatgga gaaatcacac tgacacacca ccacactatc acagggacac agacacacgg 1020
agacatcacc acatggacac actgtcacac taccacaggg acacgagaca tcacactgtc 1080
acatggacac accatcacac acatgaacac accgacacac tgccatatgg acactgccac 1140
acacactgcc acactgtcac atggacacac ctccatacca tcacaccacc acacacactg 1200
ccatgtggac acaaggacac acagacactg tcacacagat acacaaaaca ctgtcacacg 1260
gagacatcac catgcagata caccaccaca tggacatagc accagacact ctgccacaca 1320
gatacaccac cacacagaaa tgcggacaca ctgccacaca gacaccacca catcgttgcc 1380
acaatttcat gtgtcagctg gcggtgtggg cccacgact ctgggctcta atcgagaaat 1440
tacttgagaca tatagtgaag gcaaaatttt tttttatttt ctgggtaacc aagcgcgact 1500
ctgtctcaaa aaaagaaaaa aaaagcaata tactgtgtaa tcgttgacag cataattcac 1560
tattatgtag atcggagagc agaggattct gaatgcatga acatatcatt aacatttcaa 1620
tacattactc ataattactg atgaactaaa gagaaaccaa gaaatttatgg tgatagttat 1680
attgacctgg agaaatgtag acacaaaaga accgtaagat gagaaatgtg ttaacacagt 1740
ctataagggc atgcaagaat aaaaataggg gagaaaacag gagagttttt caagagcttt 1800
ctggctcatgt aagtcaactt gtatcggtta atttttaaaa ggtttattta catgcaataa 1860
actgcacata cttcaattgt acatttttgg aattcttggc attttagct ctataaaacc 1920
agcaacatat taaaatagca aacatatcca ttacctttac caccaaagtt ttcttgtgtt 1980
ttttctactc acttttttct gcctatcccc cactctcttc cacaggtaac cactgatcca 2040
cttcacgtca cttccatga gtttttattt ccaaatacat gaaatcatat gaatttctgg 2100
tttttctgtt ttgagcccaa ggagcaaggg cagaatgagg aacatgatgt ttcttwcga 2160
cagttactca tgacgtctcc atccaggact gaggggggca tccttctcca tctaggactg 2220
ggggcatcct tctccatcca gtattggggg tcctccttct ccatccagta ttgggggtca 2280

```



```

tcctctcca tccaggacct gaggggtgtc cttttctgcg cttccttgga tggcagtctt 2340
tcccttcatg tttatagtra cttaccatta aatcactgtg ccgttttttc ctaaaataaa 2400
aaaaaaaaaa aaaa 2414

```

<210> 476

<211> 3434

<212> DNA

<213> Homo sapiens

<400> 476

```

ctgtgctgca aatggggcca tatagaggaa aggagcagct ggctctggag catggtgtgc 60
actccctttg ggccttcagt ccatgtctca tgggtcgtat gacactgcgg gcttggttgg 120
tgccaagagg cagaccacag gtcactctga ggaggacttt atgttccagt ccagaaagca 180
gccagtggta ccaccaggg gacttggtgt tctgtggccc aggccagacg tagaatttga 240
caaagtcagg acggtctcag tcagagcagc atgtcgggtcc ccggggcctg tgcattgccg 300
gcagggccag gctggcttaa ggagcaagca gccacctctg ttagggggtg gcctggagca 360
ggtggagcag ccaccaacct cacgcactga aagaagcagg gatggccagg ttccaacatc 420
ctgagtggct gccacctgat ggctgatgga gcagaggcct gagggaaaagc agatggcact 480
gctttgtagt gctgttcttt gtctctcttg atctttttca gttaatgtct gttttatcag 540
agactaggat tgcaaacctt gctctttttt gctttccatt tgcttggtaa atattcctcc 600
atccctttat ttttaagccta tgtgtgtctt tgcacatgag atgggtctcc tgaatacagg 660
acaacaatgg gtctttactc tttatccaac ttgccagtct gtgtctttta actggggcat 720
ttagcccatt tacatttaag tttagtattt gttacatgtg aaatttatcc tgtcatgatg 780
ttgctagctt tttatttttc ccattagttt gcagtttctt tatagtgtca atgggtctta 840
caattcgata tgtttttgta gtggctggta ctgggtttttc ctttctacgt ttagtgtctc 900
cttcaggagc tcttgtaaca caagaatgtg gatttatttc ttgtaaggta aatatgtgga 960
tttattctgg gactgtattc tatggccttt accccaagaa tcattacttt ttaaaatgca 1020
attcaaatga gcataaaaca tttacagcct atggaaaaggc ttgtggcatt agaatcctta 1080
tttataggat tattttgtgt ttttttgaga tatggtcttt gtcatcgagg cagaagtgcc 1140
gtggtttgat cataattcac cacagccctg aactcttgag tccaagccat ccttttgcc 1200
taatctccca accagttgga tctacaagca taaggcatca tgcgtggcta attttttcac 1260
gttttttttt tttttgtcga gattatggta tcaactgtgt gctctggctg atctcaaatg 1320
tttgacctca agggatcttt ctgccacagc ctctaaagt gctaggatta tatgcattat 1380
acaccatgcc tattgtagag tattacatta ttttcaaagt cttattgtaa gagccattta 1440
ttgcctttgg cctaataaac tcaatataat atctctgaaa cttttttttg acaaattttg 1500
gggctgtatg atgagagaag ggggtttgaa actttctaata aagagttaac ttagagccat 1560
ttaagaaagg aaaaaacaca aattatcaga aaaacaacag taagatcaag tgcaaaagtt 1620
ctgtggcaaa gatgatgaga gtaagaata tatgtttgtg actcatggtg gcttttactt 1680
tgttcttgaa tttctgagta cgggttaaca tttaaagaat ctacattata gataacattt 1740
tattgcaagt aaatgtattt caaaatttgt tattggtttt gtatgagatt attctcagcc 1800
tacttcatta tcaagctata ttattttatt aatgtagttc gatgatctta cagcaaagct 1860
gaaagctgta tcttcaaaat atgtctattt gactaaaaag ttattcaaca ggagttatta 1920
tctataaaaa aatacaacag gaatataaaa aacttgagga taaaaagatg ttggaaaaag 1980
taatattaaa tcttaaaaaa catatggaaa ctacacaatg gtgaagacac attggtgaag 2040
tacaaaaata taaattggat ctagaagaaa gggcaatgca ggcaatagaa aaattagtag 2100
aaatcccttt aaaggttagt ttgtaaaatc aggttaagttt atttataatt tgctttcatt 2160
tatttcaactg caaattatat tttggatatg tatatatatt gtgcttcctc tgctgtctt 2220
acagcaattt gccttgcaga gttctaggaa aaaggtggca tgtgttttta ctttcaaaat 2280
atttaaattt ccattcattat aacaaaatca atttttcaga gtaatgattc tcaactgtgga 2340
gtcatttgat tattaagacc cgttggcata agattacatc ctctgactat aaaaatcctg 2400
gaagaaaacc taggaaatat tcgtctggac attgcacttg gcaatgaatt tatgggcgct 2460
ttggaatcct gcagataata taatgataat taaacaaaac actcagagaa actgccaacc 2520
ctaggatgaa gtatattgtt actgtgcttt gggattaaaa taagtaacta cagtttatag 2580
aactttttata ctgatacaca gacactaaaa agggaaaggg tttagatgag aagctctgct 2640

```

<213> Homo sapiens

<400> 478

Met Tyr Arg His Thr Glu Thr Leu Pro His Gly Asp Thr Val Thr Gln
 5 10 15

Ser His Gly His Thr Gly Ile Val Thr Trp Thr Asp Thr Gln Thr Tyr
 20 25 30

Gly Glu Ile Thr Trp Thr His His His Thr Ile Thr Gly Thr Gln Thr
 35 40 45

His Gly Asp Ile Thr Thr Trp Thr His Cys His Thr Thr Thr Gly Thr
 50 55 60

Arg Asp Ile Thr Leu Ser His Gly His Thr Ile Thr His Met Asn Thr
 65 70 75 80

Pro Thr His Cys His Met Asp Thr Gly Thr His Thr Ala Thr Leu Ser
 85 90 95

His Gly His Thr Ser Thr Pro Ser His His His Thr His Cys Leu Trp
 100 105 110

Thr Gln Gly His Thr Asp Thr Val Thr Gln Ile His Lys Thr Leu Ser
 115 120 125

His Gly Asp Ile Thr Met Gln Ile His His His Ser Gly Ala Val
 130 135 140

<210> 479

<211> 222

<212> PRT

<213> Homo sapiens

<400> 479

Met Tyr Arg His Thr Glu Thr Leu Pro His Gly Asp Thr Val Thr Gln
 5 10 15

Ser His Glu His Thr Gly Ile Val Thr Trp Thr Asp Thr Gln Thr Tyr
 20 25 30

Gly Glu Ile Thr Leu Thr His His His Thr Ile Thr Gly Thr Gln Thr
 35 40 45

His Gly Asp Ile Thr Thr Trp Thr His Cys His Thr Thr Thr Gly Thr
 50 55 60

Arg Asp Ile Thr Leu Ser His Gly His Thr Ile Thr His Met Asn Thr
 65 70 75 80

Pro Thr His Cys His Met Asp Thr Ala Thr His Thr Ala Thr Leu Ser
 85 90 95

His Gly His Thr Ser Ile Pro Ser His His His Thr His Cys His Val

009060"023950

Ala Gly Val Ala Arg Phe Pro Arg Pro Glu Trp Val Pro Pro Asn Gly
 130 135 140

<210> 481

<211> 167

<212> PRT

<213> Homo sapiens

<400> 481

Met His Gly Pro Gln Val Leu Ala Arg Cys Ser Glu Cys Ala Cys Pro
 5 10 15

Ala Leu Ala Ala Thr Ser Ala Gly Val Arg Leu Glu Gly Val Asp Arg
 20 25 30

Pro Pro Thr Leu Pro Ser Gln Gly Ser Gly Trp Pro Cys Ser His Ser
 35 40 45

Leu Ser Gly Cys His Leu Met Ala Asp Gly Ala Lys Ala Leu Gly Lys
 50 55 60

Ala Asp Gly Pro Trp Pro Tyr Leu Phe Val Arg Arg Thr Asp Val Pro
 65 70 75 80

Cys Pro Ala Ala Ser Glu Val Gly Gly Cys Ala Pro Ser Ser Trp Arg
 85 90 95

Ala Leu Ala Glu Val Thr Gly Cys Ser Leu Gly Pro Leu Gly Leu Ala
 100 105 110

Gln His Ala Gln Ala Ser Val Leu Leu Leu Cys Tyr Lys Trp Ser His
 115 120 125

Ile Gly Glu Thr Ser Ser His Leu Arg Ser Lys Val Tyr Ala Ala Phe
 130 135 140

Gly Gly Ser Ser Pro Cys Leu Lys Gly Leu Met Ser Leu Trp Ala Ser
 145 150 155 160

Trp Leu Ser Arg Gly Arg Pro
 165

<210> 482

<211> 143

<212> PRT

<213> Homo sapiens

<400> 482

Met Glu Pro Tyr Arg Gly Asn Lys Lys Gln Val Gln Glu Lys Gly Val

5 10 15
 Pro Cys Leu Trp Gly Ser Ser Pro Cys Leu Arg Cys His Met Ala Leu
 20 25 30
 Arg Ala Ser Trp Leu Pro Gly Gly Gly Pro Gln Ala Ile Leu Gly Arg
 35 40 45
 Thr Leu Cys Ser Ser Ala Glu Ser Ser Gln Asp Cys His Pro Gly Gly
 50 55 60
 Pro Ser Ile Ala Leu Ala Lys Pro Cys Arg Gly Val Trp Leu Leu Phe
 65 70 75 80
 Glu Pro Ala Trp Pro Pro Trp His Ala Arg Ala Pro Gly Ala Gly Thr
 85 90 95
 Leu Leu Arg Val Cys Leu Ser Cys Leu Gly Cys His Leu Cys Gly Gly
 100 105 110
 Ala Ser Gly Gly Gly Gly Pro Ala Thr Asn Leu Thr Gln Ser Arg Lys
 115 120 125
 Trp Met Ala Met Phe Pro Gln Pro Glu Trp Leu Pro Pro Asp Gly
 130 135 140

 <210> 483
 <211> 143
 <212> PRT
 <213> Homo sapiens

 <400> 483
 Met Glu Thr Gln Arg Gly Asn Lys Gln Arg Ala Gln Glu Gln Gly Val
 5 10 15
 Cys Cys Leu Trp Gly Ser Ser Pro Cys Leu Gly Ser Tyr Gly Thr Ala
 20 25 30
 Gly Phe Leu Val Ala Lys Arg Arg Thr Thr Gly Leu Leu Glu Glu Asp
 35 40 45
 Phe Thr Phe Lys Cys Arg Lys Gln Pro Lys Leu Pro Ser Met Arg Leu
 50 55 60
 Ser Leu Leu Trp Pro Trp Arg Asp Leu Lys Phe Val Pro Arg Gln Asp
 65 70 75 80
 Lys Leu Thr Arg Ser Ser Val Ser Val Ala Gly Ala Tyr Ala Cys Arg
 85 90 95
 Ala Gly Pro Gly Trp Leu Lys Glu Gln Pro Ala Thr Ser Ala Arg Val
 100 105 110

009060" 66225960

Arg Leu Val Gln Ala Glu His Pro Pro Pro His Pro Leu Glu Glu Val
 115 120 125

Gly Met Ala Arg Phe Pro Gln Pro Glu Cys Leu Pro Pro Tyr Cys
 130 135 140

<210> 484
 <211> 30
 <212> PRT
 <213> Homo Sapien

<400> 484
 Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
 1 5 10 15
 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile
 20 25 30

<210> 485
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 485
 gggaagctta tcacctatgt gccgcctctg c 31

<210> 486
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 486
 gcgaattctc acgctgagta tttggcc 27

<210> 487
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 487
 cccgaattct tagctgccca tccgaacgcc ttcac 36

<210> 488
 <211> 33
 <212> DNA
 <213> Artificial Sequence

009060"6225960

<223> Made in a lab

gggaagcttc ttccccggct gcaccagctg tgc

33

<211> 19

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala

1 5 10 15

Ser Val Ala

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys

1	5	10	15
---	---	----	----

Leu Ser His Ser

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu

1	5	10	15
---	---	----	----

Thr Gly Phe Thr

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Ala Leu Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr
1 5 10 15
Leu Ala Ser Leu
20

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Tyr Thr Leu Ala Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro
1 5 10 15
Lys Tyr Arg Gly
20

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser
1 5 10 15
Leu Met Ile Ser
20

<211> 20

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro Lys Pro Gly Ala Pro
1 5 10 15
Phe Pro Asn Gly
20

<211> 21

<212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 496
 Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu Leu
 1 5 10 15
 Pro Pro Pro Pro Ala
 20

<210> 497
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 497
 Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val
 1 5 10 15
 Ser Val Arg Val
 20

<210> 498
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 498
 Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val
 1 5 10 15
 Val Pro Gly Arg
 20

<210> 499
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 499
 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 1 5 10 15
 Ser Ala Phe Leu
 20

009060"624960

<210> 500
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 500
 Leu Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met
 1 5 10 15
 Gly Ser Ile Val
 20

<210> 501
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 501
 Phe Met Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met
 1 5 10 15
 Val Ser Ala Ala
 20

<210> 502
 <211> 414
 <212> DNA
 <213> Homo Sapien

<220>
 <221> misc_feature
 <222> (1)...(414)
 <223> n=A,T,C or G

<400> 502
 caccatggag acaggcctgc gctggctttt cctggctcgt gtgctcaaag gtgtccaatg 60
 tcagtcggtg gaggagtccg ggggtcgcct ggtcacgcct gggacacctt tgacantcac 120
 ctgtagagtt tttggaatng acctcagtag caatgcaatg agctgggtcc gccaggctcc 180
 agggaagggg ctggaatgga tcggagccat tgataattgt ccacantacg cgacctgggc 240
 gaaaggccga ttnatnattn ccaaaacctn gaccacggtg gatttgaaaa tgaccagtcc 300
 gacaaccgag gacacggcca cctatntttg tggcagaatg aatactggta atagtgggtg 360
 gaagaatatt tgggggccag gcaccctggt caccgtntcc tcagggaac ctaa 414

<210> 503
 <211> 379
 <212> DNA
 <213> Homo Sapien

<220>
 <221> misc_feature

<222> (1)...(379)

<223> n=A,T,C or G

<400> 503

atnccgatggg	gcttgggtcaa	aggtgtccag	tgctcagtcgg	tggaggagtc	cggggggtcgc	60
ctgggtcacgc	ctgggacacc	cctgacactc	acctgcaccg	tntctggatt	ngacatcagt	120
agctatggag	tgagctgggt	ccgccaggct	ccagggaagg	ggctgggnata	catcggatca	180
ttagtagtag	tggtacattt	tacgcgagct	gggcgaaagg	ccgattcacc	atttccaaaa	240
cctngaccac	ggtggatttg	aaaatcacca	gtttgacaac	cgaggacacg	gccacctatt	300
tntgtgccag	aggggggttt	aattataaag	acatttgggg	cccaggcacc	ctgggtcaccg	360
tntccttagg	gcaacctaa					379

<210> 504

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 504

Gly	Phe	Thr	Asn	Tyr	Thr	Asp	Phe	Glu	Asp	Ser	Pro	Tyr	Phe	Lys	Glu
1				5				10						15	
Asn	Ser	Ala													

<210> 505

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 505

Lys	Glu	Asn	Ser	Ala	Phe	Pro	Pro	Phe	Cys	Cys	Asn	Asp	Asn	Val	Thr
1				5				10						15	
Asn	Thr	Ala	Asn												
				20											

<210> 506

<211> 407

<212> DNA

<213> Homo Sapien

<400> 506

atggagacag	gcctgcgctg	gcttctcctg	gtcgtgcgc	tcaaagggtg	ccagtgtcag	60
tgcctggagg	agtcggggg	tcgcctggtc	acgcctggga	cacccctgac	actcacctgc	120
aacgtctctg	gattctccct	cagtagcaat	gcaatgatct	gggtccgcca	ggctccaggg	180
aaggggctgg	aatacatcgg	atacattagt	tatggtggtg	gcgcatacta	cgcgagctgg	240
gtgaaaggcc	gattcaccat	ctccaaaacc	tcgaccacgg	tggatctgag	aatgaccagt	300
ctgacaaccg	aggacacggc	cacctatttc	tgtgccagaa	atagtgattt	tagtggtatg	360
ttgtggggcc	caggcacctt	ggtcaccgtc	tcctcagggc	aacctaa		407

009960 "6446960

<400> 510
 Pro Glu Tyr Asn Arg Pro Leu Leu Ala Asn Asp Leu Met Leu Ile
 1 5 10 15

<210> 511
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 511

Tyr His Pro Ser Met Phe Cys Ala Gly Gly Gly Gln Asp Gln Lys
 1 5 10 15

<210> 512
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 512
 Asp Ser Gly Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu
 1 5 10 15

<210> 513
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 513
 Ala Pro Cys Gly Gln Val Gly Val Pro Asx Val Tyr Thr Asn Leu
 1 5 10 15

<210> 514
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 514
 Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
 1 5 10 15

009050"626360

<400> 519

<210> 520

<211> 25

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Made in a lab

<400> 520

Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr
1 5 10 15
Glu Ala Arg Arg His Tyr Asp Glu Gly
20 25

<210> 521

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 521

Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu Leu
1 5 10 15
Pro Pro Pro Pro Ala
20

<210> 522

<211> 20

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Made in a lab

<400> 522

Leu Leu Val Val Pro Ala Ile Lys Lys Asp Tyr Gly Ser Gln Glu Asp
1 5 10 15
Phe Thr Gln Val
20

<210> 523

<211> 254

<212> PRT

<213> Artificial Sequence

<223> Made in a lab

<221> VARIANT

<222> (1) ... (254)

<223> Xaa = any amino acid

<400> 523

Met	Ala	Thr	Ala	Gly	Asn	Pro	Trp	Gly	Trp	Phe	Leu	Gly	Tyr	Leu	Ile
1				5				10						15	
Leu	Gly	Val	Ala	Gly	Ser	Leu	Val	Ser	Gly	Ser	Cys	Ser	Gln	Ile	Ile
			20					25					30		
Asn	Gly	Glu	Asp	Cys	Ser	Pro	His	Ser	Gln	Pro	Trp	Gln	Ala	Ala	Leu
		35					40					45			
Val	Met	Glu	Asn	Glu	Leu	Phe	Cys	Ser	Gly	Val	Leu	Val	His	Pro	Gln
	50					55					60				
Trp	Val	Leu	Ser	Ala	Thr	His	Cys	Phe	Gln	Asn	Ser	Tyr	Thr	Ile	Gly
65					70					75				80	
Leu	Gly	Leu	His	Ser	Leu	Glu	Ala	Asp	Gln	Glu	Pro	Gly	Ser	Gln	Met
				85					90					95	
Val	Glu	Ala	Ser	Leu	Ser	Val	Arg	His	Pro	Glu	Tyr	Asn	Arg	Pro	Leu
			100					105					110		
Leu	Ala	Asn	Asp	Leu	Met	Leu	Ile	Lys	Leu	Asp	Glu	Ser	Val	Ser	Glu
		115					120					125			
Ser	Asp	Thr	Ile	Arg	Ser	Ile	Ser	Ile	Ala	Ser	Gln	Cys	Pro	Thr	Ala
	130					135					140				
Gly	Asn	Ser	Cys	Leu	Val	Ser	Gly	Trp	Gly	Leu	Leu	Ala	Asn	Gly	Arg
145				150						155				160	
Met	Pro	Thr	Val	Leu	Gln	Cys	Val	Asn	Val	Ser	Val	Val	Ser	Glu	Glu
				165					170					175	
Val	Cys	Ser	Lys	Leu	Tyr	Asp	Pro	Leu	Tyr	His	Pro	Ser	Met	Phe	Cys
			180					185					190		
Ala	Gly	Gly	Gly	Gln	Xaa	Gln	Xaa	Asp	Ser	Cys	Asn	Gly	Asp	Ser	Gly
		195					200					205			
Gly	Pro	Leu	Ile	Cys	Asn	Gly	Tyr	Leu	Gln	Gly	Leu	Val	Ser	Phe	Gly
	210					215					220				
Lys	Ala	Pro	Cys	Gly	Gln	Val	Gly	Val	Pro	Gly	Val	Tyr	Thr	Asn	Leu
225				230						235				240	
Cys	Lys	Phe	Thr	Glu	Trp	Ile	Glu	Lys	Thr	Val	Gln	Ala	Ser		
				245					250						

<210> 524

<211> 765

<212> DNA

<213> Homo sapien

<400> 524

atggccacag	caggaaatcc	ctggggctgg	tctctggggt	acctcatcct	tgggtgcgca	60
ggatcgctcg	tctctggtag	ctgcagccaa	atcataaacg	gcgaggactg	cagcccgcac	120
tcgcagccct	ggcaggcggc	actggtcatg	gaaaacgaat	tgttctgctc	gggcgtcctg	180
gtgcatccgc	agtgggtgct	gtcagccgca	cactgtttcc	agaactccta	caccatcggt	240
ctgggcctgc	acaqtcttga	ggccgaccaa	gagccaggga	gccagatggg	ggaggccagc	300


```

atgagttcct gcaacttcac acatgccacc tttgtgctta ttggtatccc aggattagag 60
aaagcccatt tctgggttgg cttccccctc ctttccatgt atgtagtggc aatgttttga 120
aactgcatcg tgggtcttcat cgtaaggacg gaacgcagcc tgcacgctcc gatgtacctc 180
tttctctgca tgcttgcagc cattgacctg gccttatcca catccaccat gcctaagatc 240
cttgcccttt tctgggttga ttcccgagag attagctttg aggcctgtct taccagatg 300
ttctttattc atgccctctc agccattgaa tccaccatcc tgctggccat ggcctttgac 360
cgttatgtgg ccatctgcca cccactgcgc catgctgcag tgctcaacaa tacagtaaca 420
gcccagattg gcatcgtggc tgtggtcgcg ggatccctct tttttttccc actgcctctg 480
ctgatcaagc ggctggcctt ctgccactcc aatgtcctct cgcactccta ttgtgtccac 540
caggatgtaa tgaagttggc ctatgcagac actttgcccc atgtgggtata tgggtcttact 600
gccattctgc tgggtcatggg cgtggacgta atgttcatct ccttgtccta ttttctgata 660
atacgaacgg ttctgcaact gccttccaag tcagagcggg ccaaggcctt tggaacctgt 720
gtgtcacaca ttggtgtggt actcgccttc tatgtgccac ttattggcct ctcaagtgtgta 780
caccgctttg gaaacagcct tcatcccatg gtgcgtgttg tcatgggtga catctacctg 840
ctgctgcttc ctgtcatcaa tcccatcatc tatggtgcc aacacaaaca gatcagaaca 900
cgggtgctgg ctatgttcaa gatcagctgt gacaaggact tgcaggctgt gggaggcaag 960
tga

```

<210> 527

<211> 320

<212> PRT

<213> Homo sapiens

<400> 527

Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly Ile
5 10 15

Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu Leu Ser
20 25 30

Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val Phe Ile Val
35 40 45

Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu Phe Leu Cys Met
50 55 60

Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser Thr Met Pro Lys Ile
65 70 75 80

Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu Ile Ser Phe Glu Ala Cys
85 90 95

Leu Thr Gln Met Phe Phe Ile His Ala Leu Ser Ala Ile Glu Ser Thr
100 105 110

Ile Leu Leu Ala Met Ala Phe Asp Arg Tyr Val Ala Ile Cys His Pro
115 120 125

Leu Arg His Ala Ala Val Leu Asn Asn Thr Val Thr Ala Gln Ile Gly
130 135 140

Ile Val Ala Val Val Arg Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu
145 150 155 160

000050"626960

```

<210> 528
<211> 20
<212> DNA
<213> Homo Sapien

<400> 528
actatgggtcc agaggctgtg
20

<210> 529
<211> 20
<212> DNA
<213> Homo Sapien

<400> 529
atcacctatg tgccgcctct
20

<210> 530
<211> 1852
<212> DNA
<213> Homo sapiens

<400> 530
ggcacgagaa ttaaaaccct cagcaaaaaca ggcatagaag ggacatacct taaagtaata 60
aaaaccacct atgacaagcc cacagccaac ataatactaa atgggggaaaa gttagaagca 120

```


<212> PRT

<213> Homo sapiens

<400> 532

```

Met His Leu Ser Phe Pro Ala Phe Leu Pro Pro Trp Met Asp Arg Gly
              5              10              15

Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp His Asn Asp Ser Ser
              20              25              30

Val Lys Thr Leu Gly Ser Lys Arg Cys Lys Trp Cys Cys His Cys Phe
              35              40              45

Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val Val Ala Trp Gly Asp
              50              55              60

Tyr Asp Asp Ser Ala Phe Met Asp Pro Arg Tyr His Val His Gly Glu
              65              70              75              80

Asp Leu Asp Lys Leu His Arg Ala Ala Trp Trp Gly Lys Val Pro Arg
              85              90              95

Lys Asp Leu Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys Arg Asp
              100             105             110

Lys Gln Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser
              115             120             125

Glu Val Val Lys Leu Val Leu Asp Arg Arg Cys Gln Leu Asn Val Leu
              130             135             140

Asp Asn Lys Lys Arg Thr Ala Leu Thr Lys Ala Val Gln Cys Gln Glu
              145             150             155             160

Asp Glu Cys Ala Leu Met Leu Leu Glu His Gly Thr Asp Pro Asn Ile
              165             170             175

Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr Ala Val Tyr Asn Glu
              180             185             190

Asp Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu
              195             200             205

Ser Lys Asn Lys His Gly Leu Thr Pro Leu Leu Leu Gly Ile His Glu
              210             215             220

Gln Lys Gln Gln Val Val Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu
              225             230             235             240

Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu Ile Leu Ala Val Cys
              245             250             255

Cys Gly Ser Ala Ser Ile Val Ser Pro Leu Leu Glu Gln Asn Val Asp
              260             265             270

```

009050"6225000

Val Ser Ser Gln Asp Leu Glu Arg Arg Pro Glu Ser Met Leu Phe Leu
 275 280 285

Val Ile Ile Met
 290

<210> 533
 <211> 801
 <212> DNA
 <213> Homo sapiens

<400> 533
 atgtacaagc ttccagtgcaa caactgtgct acaaattggag ccacagagag gaaacaagca 60
 gcaggctcag gagcagggtta tgcgctgcct tcggctctcc aatccatgcc tcagggtccc 120
 tatgccactg cacgattctt gggtgccaag aggccaaacca caggccatct tgagaaggag 180
 tttatgttcc actgcagaaa gcagccagga tcaccatcca ggggacttgg tcttctgtgg 240
 ccctggccag acatagaatt tgtgccaagg caggacaagc tcaatcagag cagcgtgtta 300
 gtacctcaaa tctgtgctg ccagacaagg ccaaactggc tcaatgagca accagccacc 360
 tctgcagggg tgcgtctgga ggaggtggac cagccaccaa ccttaccag tcaaggaagt 420
 ggatggccat gttccacag cctgagtggc tgccacctga tggctgatat agcaaaggcc 480
 ttaggaaaag cagatggccc ttggccctac ctttttgta gaagaactga tgttccatgt 540
 cctgcagcga gtgaggttg tggtctgtgc ccagctcct ggcacaccct cgcagagggtg 600
 actggttgct ctttgagccc tcttagcctt gccagcatg cacaagcctc agtgctacta 660
 ctgtgctaca aatggagcca tataggggaa acgagcagcc atctcaggag caagggtgat 720
 gctgcctttg ggggctccag tccttgctc aaggggtotta tgtcactgtg ggcttcttgg 780
 ttgccaagag gcagaccata g 801

<210> 534
 <211> 266
 <212> PRT
 <213> Homo sapiens

<400> 534
 Met Tyr Lys Leu Gln Cys Asn Asn Cys Ala Thr Asn Gly Ala Thr Glu
 5 10 15
 Arg Lys Gln Ala Ala Gly Ser Gly Ala Gly Tyr Ala Leu Pro Ser Ala
 20 25 30
 Leu Gln Ser Met Pro Gln Gly Ser Tyr Ala Thr Ala Arg Phe Leu Val
 35 40 45
 Ala Lys Arg Pro Thr Thr Gly His Leu Glu Lys Glu Phe Met Phe His
 50 55 60
 Cys Arg Lys Gln Pro Gly Ser Pro Ser Arg Gly Leu Gly Leu Leu Trp
 65 70 75 80
 Pro Trp Pro Asp Ile Glu Phe Val Pro Arg Gln Asp Lys Leu Thr Gln
 85 90 95
 Ser Ser Val Leu Val Pro Gln Ile Cys Ala Cys Gln Thr Arg Pro Asn

100 105 110
 Trp Leu Asn Glu Gln Pro Ala Thr Ser Ala Gly Val Arg Leu Glu Glu
 115 120 125
 Val Asp Gln Pro Pro Thr Leu Pro Ser Gln Gly Ser Gly Trp Pro Cys
 130 135 140
 Ser His Ser Leu Ser Gly Cys His Leu Met Ala Asp Ile Ala Lys Ala
 145 150 155 160
 Leu Gly Lys Ala Asp Gly Pro Trp Pro Tyr Leu Phe Val Arg Arg Thr
 165 170 175
 Asp Val Pro Cys Pro Ala Ala Ser Glu Val Gly Gly Cys Ala Pro Ser
 180 185 190
 Ser Trp His Thr Leu Ala Glu Val Thr Gly Cys Ser Leu Ser Pro Leu
 195 200 205
 Ser Leu Ala Gln His Ala Gln Ala Ser Val Leu Leu Leu Cys Tyr Lys
 210 215 220
 Trp Ser His Ile Gly Glu Thr Ser Ser His Leu Arg Ser Lys Val Tyr
 225 230 235 240
 Ala Ala Phe Gly Gly Ser Ser Pro Cys Leu Lys Gly Leu Met Ser Leu
 245 250 255
 Trp Ala Ser Trp Leu Pro Arg Gly Arg Pro
 260 265

<210> 535

<211> 6082

<212> DNA

<213> Homo sapiens

<400> 535

cctccactat tacagcttat aggaaattac aatccacttt acaggcctca aagggttcatt 60
 ctggccgagc ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag 120
 cggagcccg cggccactgcc gcctgatcag cgcgaccccg gcccgcgccc gccccgcccc 180
 gcaagatgct gcccggtgtac caggaggtga agcccaaccc gctgcaggac gcgaacctct 240
 gctcacgcgt gttcttcttg tggctcaatc ccttgtttaa aattggccat aaacggagat 300
 tagaggaaga tgatatgtat tcagtgtctc cagaagaccg ctacacgcac cttggagagg 360
 agttgcaagg gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt 420
 ctttaacaag agcaatcata aagtgttact ggaaatctta tttagttttg ggaattttta 480
 cgtaattga ggaaagtgcc aaagtaatcc agcccatatt tttgggaaaa attattaatt 540
 attttgaaaa ttatgatccc atggattctg tggctttgaa cacagcgtac gcctatgcca 600
 cgggtgctgac tttttgcacg ctcatcttgg ctatactgca tcacttatat ttttatcacg 660
 ttcagtgtgc tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac 720
 ttcgtcttag taacatggcc atggggaaga caaccacagg ccagatagtc aatctgctgt 780
 ccaatgatgt gaacaagttt gatcaggtga cagtgttctt acacttcctg tgggcaggac 840
 cactgcaggc gatcgagtg actgcctac tctggatgga gataggaata tcgtgccttg 900

325 330 335
 Thr Thr Tyr Val Leu Leu Gly Ser Val Ile Thr Ala Ser Arg Val Phe
 340 345 350
 Val Ala Val Thr Leu Tyr Gly Ala Val Arg Leu Thr Val Thr Leu Phe
 355 360 365
 Phe Pro Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg
 370 375 380
 Arg Ile Gln Thr Phe Leu Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg
 385 390 395 400
 Gln Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr
 405 410 415
 Ala Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser
 420 425 430
 Phe Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly
 435 440 445
 Ala Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro
 450 455 460
 Ser His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
 465 470 475 480
 Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly
 485 490 495
 Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala
 500 505 510
 Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile
 515 520 525
 Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn
 530 535 540
 Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp
 545 550 555 560
 Pro Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu
 565 570 575
 Cys Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His
 580 585 590
 Gln Leu Gln Tyr Leu Lys Ala Ala Ser Gln Ile Leu Ile Leu Lys Asp
 595 600 605
 Gly Lys Met Val Gln Lys Gly Thr Tyr Thr Glu Phe Leu Lys Ser Gly

000000"000000

900	905	910
Thr Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp		
915	920	925
Ala His Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr		
930	935	940
Ser Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val		
945	950	955
Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp Ala		
	965	970
		975
Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met Gly Met		
	980	985
		990
Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile		
	995	1000
		1005
Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro		
1010	1015	1020
Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val		
1025	1030	1035
		1040
Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro Leu		
	1045	1050
		1055
Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys Val Gly		
	1060	1065
		1070
Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile Ser Ala Leu		
	1075	1080
		1085
Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp Lys Ile Leu		
1090	1095	1100
Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met Ser Ile Ile		
1105	1110	1115
		1120
Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp		
	1125	1130
		1135
Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu		
	1140	1145
		1150
Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr		
	1155	1160
		1165
Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu		
1170	1175	1180
Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile		

1185 1190 1195 1200
 Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln
 1205 1210 1215

 Lys Lys Ser Gly Arg Asn Leu Pro Thr Ala Pro Cys
 1220 1225

 <210> 538
 <211> 1261
 <212> PRT
 <213> Homo sapiens

 <400> 538
 Met Tyr Ser Val Leu Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu
 5 10 15

 Leu Gln Gly Phe Trp Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala
 20 25 30

 Gln Lys Pro Ser Leu Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys Ser
 35 40 45

 Tyr Leu Val Leu Gly Ile Phe Thr Leu Ile Glu Glu Ser Ala Lys Val
 50 55 60

 Ile Gln Pro Ile Phe Leu Gly Lys Ile Ile Asn Tyr Phe Glu Asn Tyr
 65 70 75 80

 Asp Pro Met Asp Ser Val Ala Leu Asn Thr Ala Tyr Ala Tyr Ala Thr
 85 90 95

 Val Leu Thr Phe Cys Thr Leu Ile Leu Ala Ile Leu His His Leu Tyr
 100 105 110

 Phe Tyr His Val Gln Cys Ala Gly Met Arg Leu Arg Val Ala Met Cys
 115 120 125

 His Met Ile Tyr Arg Lys Ala Leu Arg Leu Ser Asn Met Ala Met Gly
 130 135 140

 Lys Thr Thr Thr Gly Gln Ile Val Asn Leu Leu Ser Asn Asp Val Asn
 145 150 155 160

 Lys Phe Asp Gln Val Thr Val Phe Leu His Phe Leu Trp Ala Gly Pro
 165 170 175

 Leu Gln Ala Ile Ala Val Thr Ala Leu Leu Trp Met Glu Ile Gly Ile
 180 185 190

 Ser Cys Leu Ala Gly Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln
 195 200 205

 Ser Cys Phe Gly Lys Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr

	500					505					510				
Tyr	Leu	Leu	Asp	Asp	Pro	Leu	Ser	Ala	Val	Asp	Ala	Glu	Val	Ser	Arg
	515					520					525				
His	Leu	Phe	Glu	Leu	Cys	Ile	Cys	Gln	Ile	Leu	His	Glu	Lys	Ile	Thr
	530					535					540				
Ile	Leu	Val	Thr	His	Gln	Leu	Gln	Tyr	Leu	Lys	Ala	Ala	Ser	Gln	Ile
545	550					555					560				
Leu	Ile	Leu	Lys	Asp	Gly	Lys	Met	Val	Gln	Lys	Gly	Thr	Tyr	Thr	Glu
	565					570					575				
Phe	Leu	Lys	Ser	Gly	Ile	Asp	Phe	Gly	Ser	Leu	Leu	Lys	Lys	Asp	Asn
	580					585					590				
Glu	Glu	Ser	Glu	Gln	Pro	Pro	Val	Pro	Gly	Thr	Pro	Thr	Leu	Arg	Asn
	595					600					605				
Arg	Thr	Phe	Ser	Glu	Ser	Ser	Val	Trp	Ser	Gln	Gln	Ser	Ser	Arg	Pro
	610					615					620				
Ser	Leu	Lys	Asp	Gly	Ala	Leu	Glu	Ser	Gln	Asp	Thr	Glu	Asn	Val	Pro
625	630					635					640				
Val	Thr	Leu	Ser	Glu	Glu	Asn	Arg	Ser	Glu	Gly	Lys	Val	Gly	Phe	Gln
	645					650					655				
Ala	Tyr	Lys	Asn	Tyr	Phe	Arg	Ala	Gly	Ala	His	Trp	Ile	Val	Phe	Ile
	660					665					670				
Phe	Leu	Ile	Leu	Leu	Asn	Thr	Ala	Ala	Gln	Val	Ala	Tyr	Val	Leu	Gln
	675					680					685				
Asp	Trp	Trp	Leu	Ser	Tyr	Trp	Ala	Asn	Lys	Gln	Ser	Met	Leu	Asn	Val
	690					695					700				
Thr	Val	Asn	Gly	Gly	Gly	Asn	Val	Thr	Glu	Lys	Leu	Asp	Leu	Asn	Trp
705	710					715					720				
Tyr	Leu	Gly	Ile	Tyr	Ser	Gly	Leu	Thr	Val	Ala	Thr	Val	Leu	Phe	Gly
	725					730					735				
Ile	Ala	Arg	Ser	Leu	Leu	Val	Phe	Tyr	Val	Leu	Val	Asn	Ser	Ser	Gln
	740					745					750				
Thr	Leu	His	Asn	Lys	Met	Phe	Glu	Ser	Ile	Leu	Lys	Ala	Pro	Val	Leu
	755					760					765				
Phe	Phe	Asp	Arg	Asn	Pro	Ile	Gly	Arg	Ile	Leu	Asn	Arg	Phe	Ser	Lys
	770					775					780				
Asp	Ile	Gly	His	Leu	Asp	Asp	Leu	Leu	Pro	Leu	Thr	Phe	Leu	Asp	Phe

785					790						795				800
Ile	Gln	Thr	Leu	Leu	Gln	Val	Val	Gly	Val	Val	Ser	Val	Ala	Val	Ala
				805					810					815	
Val	Ile	Pro	Trp	Ile	Ala	Ile	Pro	Leu	Val	Pro	Leu	Gly	Ile	Ile	Phe
			820					825					830		
Ile	Phe	Leu	Arg	Arg	Tyr	Phe	Leu	Glu	Thr	Ser	Arg	Asp	Val	Lys	Arg
		835					840					845			
Leu	Glu	Ser	Thr	Thr	Arg	Ser	Pro	Val	Phe	Ser	His	Leu	Ser	Ser	Ser
		850				855					860				
Leu	Gln	Gly	Leu	Trp	Thr	Ile	Arg	Ala	Tyr	Lys	Ala	Glu	Glu	Arg	Cys
865					870					875					880
Gln	Glu	Leu	Phe	Asp	Ala	His	Gln	Asp	Leu	His	Ser	Glu	Ala	Trp	Phe
				885					890					895	
Leu	Phe	Leu	Thr	Thr	Ser	Arg	Trp	Phe	Ala	Val	Arg	Leu	Asp	Ala	Ile
			900					905					910		
Cys	Ala	Met	Phe	Val	Ile	Ile	Val	Ala	Phe	Gly	Ser	Leu	Ile	Leu	Ala
		915					920					925			
Lys	Thr	Leu	Asp	Ala	Gly	Gln	Val	Gly	Leu	Ala	Leu	Ser	Tyr	Ala	Leu
	930					935					940				
Thr	Leu	Met	Gly	Met	Phe	Gln	Trp	Cys	Val	Arg	Gln	Ser	Ala	Glu	Val
945					950					955					960
Glu	Asn	Met	Met	Ile	Ser	Val	Glu	Arg	Val	Ile	Glu	Tyr	Thr	Asp	Leu
				965					970					975	
Glu	Lys	Glu	Ala	Pro	Trp	Glu	Tyr	Gln	Lys	Arg	Pro	Pro	Pro	Ala	Trp
			980					985					990		
Pro	His	Glu	Gly	Val	Ile	Ile	Phe	Asp	Asn	Val	Asn	Phe	Met	Tyr	Ser
		995					1000					1005			
Pro	Gly	Gly	Pro	Leu	Val	Leu	Lys	His	Leu	Thr	Ala	Leu	Ile	Lys	Ser
	1010					1015					1020				
Gln	Glu	Lys	Val	Gly	Ile	Val	Gly	Arg	Thr	Gly	Ala	Gly	Lys	Ser	Ser
1025					1030					1035					1040
Leu	Ile	Ser	Ala	Leu	Phe	Arg	Leu	Ser	Glu	Pro	Glu	Gly	Lys	Ile	Trp
				1045					1050					1055	
Ile	Asp	Lys	Ile	Leu	Thr	Thr	Glu	Ile	Gly	Leu	His	Asp	Leu	Arg	Lys
			1060					1065					1070		
Lys	Met	Ser	Ile	Ile	Pro	Gln	Glu	Pro	Val	Leu	Phe	Thr	Gly	Thr	Met

<400> 540

Ala Val Val Thr Ala Ser Ala Ala Leu

1

5

<210> 541

<211> 14

<212> PRT

<213> Homo sapiens

<400> 541

Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu

5

10

<210> 542

<211> 15

<212> PRT

<213> Homo sapiens

<400> 542

Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala

5

10

15

<210> 543

<211> 12

<212> PRT

<213> Homo sapiens

<400> 543

Phe Met Gly Ser Ile Val Gln Leu Ser Gln Ser Val

5

10

<210> 544

<211> 18

<212> PRT

<213> Homo sapiens

<400> 544

Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val Glu Glu Lys Phe

5

10

15

Met Thr

<210> 545

<211> 18

<212> PRT

<213> Homo sapiens

<400> 545

000000"000000

<212> PRT

<213> Homo sapiens

<400> 549

Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg
 5 10 15

Gln Ala

<210> 550

<211> 14

<212> PRT

<213> Homo sapiens

<400> 550

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe
 5 10

<210> 551

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 551

Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala
 5 10

<210> 552

<211> 2577

<212> DNA

<213> Homo sapiens

<400> 552

```
agcatatgta acatgacctg tgcttcagtg ttcttttgtg atcaaaaatt ccttactttt 60
agtttttttat ctatggtaga accaccaga gcaggggtcc tcaactccca ggccacagac 120
tcataccagt ccacggacta ttatgaacca caccacacag gaggaggtga gcactaggca 180
agccaaggaa gcttcacctg tacttacagc cacacgccat ggctcatatt acagcctgaa 240
ctctgacctc actcagatca gtgataacat tagaaactca ttggagcacg aaccctgttg 300
tgaactgcct atccgaagga tctaggttgt gtgcttcgta tgagaatcta atgccagatg 360
atctatcatt gtctcacttt gccccagat aagaccatct agttgcagaa aaataagctc 420
agagcttcca ctgattctac attatggata tgtgccgccg aagcaagcac aaagccctac 480
ttttacacat gcctagtgat gcttcatgga caaggcttgg ctctgttgag tccaactaac 540
ctacctgaga ttctgagatt tctcttcaat ggcttcctgt gagctagagt ttgaaaatat 600
cttaaaatct tgagctagag atggaagtag cttggacgat tttcattatc atgtaaatcg 660
ggctactcaa ggggccaacc acagctggga gccactgctc aggggaaggt tcatatggga 720
ctttctactg cccaagggtt tatacaggat ataaagggtc ctcacagtat agatctggta 780
gcaaagaaga agaaacaaac actgatctct ttctgccacc cctctgaccc tttggaactc 840
ctctgaccct ttagaacaag cctaccta atctgctaga gaaaagacca acaacgcct 900
```



```
<210> 553
<211> 58
<212> PRT
<213> Homo sapiens
```

Ser Ile Cys Asn Met Thr Cys Ala Ser Val Phe Phe Cys Asp Gln Lys
5 10 15

Val Leu Asn Ser Gln Ala Thr Asp Ser Tyr Gln Ser Thr Asp Tyr Tyr
35 40 45

```
<210> 554
<211> 59
<212> PRT
<213> Homo sapiens
```

<400> 554

Leu Gln Lys Asn Lys Leu Arg Ala Ser Thr Asp Ser Thr Leu Trp Ile
 5 10 15

Cys Ala Ala Glu Ala Ser Thr Lys Pro Tyr Phe Tyr Thr Cys Leu Val
 20 25 30

Met Leu His Gly Gln Gly Leu Ala Leu Leu Ser Pro Thr Asn Leu Pro
 35 40 45

Glu Ile Leu Arg Phe Leu Phe Asn Gly Phe Leu
 50 55

<210> 555

<211> 71

<212> PRT

<213> Homo sapiens

<400> 555

Leu Gly Arg Phe Ser Leu Ser Cys Lys Ser Gly His Ser Arg Gly Gln
 5 10 15

Pro Gln Leu Gly Ala Thr Ala Gln Gly Lys Val His Met Gly Leu Ser
 20 25 30

Thr Ala Gln Gly Ser Ile Gln Asp Ile Lys Val Pro His Ser Ile Asp
 35 40 45

Leu Val Ala Lys Lys Lys Lys Gln Thr Leu Ile Ser Phe Cys His Pro
 50 55 60

Ser Asp Pro Leu Glu Leu Leu
 65 70

<210> 556

<211> 81

<212> PRT

<213> Homo sapiens

<400> 556

Asn His Pro Glu Gln Gly Ser Ser Thr Pro Arg Pro Gln Thr His Thr
 5 10 15

Ser Pro Arg Thr Ile Met Asn His Thr Thr Gln Glu Glu Val Ser Thr
 20 25 30

Arg Gln Ala Lys Glu Ala Ser Pro Val Leu Thr Ala Thr Arg His Gly
 35 40 45

Ser Tyr Tyr Ser Leu Asn Ser Ala Ser Thr Gln Ile Ser Asp Asn Ile
 50 55 60

Arg Asn Ser Leu Glu His Glu Pro Cys Cys Glu Leu Pro Ile Arg Arg

000000" 5225960

<210> 559
<211> 50

<213> Homo sapiens

Thr Leu Pro Pro Leu Arg Ser Val Ile Thr Leu Glu Thr His Trp Ser
5 10 15

Ser Tyr Glu Asn Leu Met Pro Asp Asp Leu Ser Leu Ser His Phe Ala
35 40 45

<213> Homo sapiens

Ile Gly Ser Leu Lys Gly Pro Thr Thr Ala Gly Ser His Cys Ser Gly
5 10 15

Lys Gly Ala Ser Gln Tyr Arg Ser Gly Ser Lys Glu Glu Glu Thr Asn
35 40 45

<213> Homo sapiens

<223> Xaa = Any amino acid

Val Leu His Leu Asp Gln Met Asn Asn Val Gly Ile Xaa Met Asp Lys
5 10 15

Asn Leu Ser Cys Phe Leu Ser Xaa Phe Trp Leu Met Gln Gly Thr Asn

```

35
40
45
Ser Leu Pro Arg Glu Asn Tyr Leu Asn
50 55

<210> 562
<211> 59
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(59)
<223> Xaa = Any amino acid

<400> 562
Asp Leu Tyr Pro Xaa Arg Ser Gln His Cys Ser Phe Asp Pro Ser Val
5 10 15

Ala Pro Met His Gly Ile Lys Asn Ser Ile Thr Ser Leu Ile Phe Leu
20 25 30

Ile Ser Tyr Leu Xaa Leu Glu Met Ser Ser Leu Ser Glu Ser Leu Val
35 40 45

Leu Ser Ser Gly Asp Tyr Val Leu Asp Thr Pro
50 55

<210> 563
<211> 79
<212> PRT
<213> Homo sapiens

<400> 563
Cys Phe Leu Phe Pro Tyr Leu Trp Leu Tyr Ala Gln Pro Leu Phe Pro
5 10 15

Lys Gln Gln Pro Pro Ala Leu Ala Pro Gly His Pro Asp Phe Ile His
20 25 30

Thr Gln Asn Glu Gln Ile Asp Pro Ser Pro His Ile Gln Asn Leu Met
35 40 45

Trp Asn Pro His Leu Ser Gln Glu Leu Ala Glu Thr Phe Met Val Arg
50 55 60

Asp Pro Leu Arg Pro Leu Leu Val Phe Ser Leu Ala Asp Ile Arg
65 70 75

<210> 564
<211> 64

```

<212> PRT

<213> Homo sapiens

<400> 564

Ala Cys Ser Lys Gly Ser Glu Glu Phe Gln Arg Val Arg Gly Val Ala
 5 10 15

Glu Arg Asp Gln Cys Leu Phe Leu Leu Leu Cys Tyr Gln Ile Tyr Thr
 20 25 30

Val Arg His Leu Tyr Ile Leu Tyr Arg Thr Leu Gly Ser Arg Lys Ser
 35 40 45

His Met Asn Leu Pro Leu Ser Ser Gly Ser Gln Leu Trp Leu Ala Pro
 50 55 60

<210> 565

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> (1)...(57)

<223> Xaa = Any amino acid

<400> 565

Leu Tyr Tyr Cys Ser Tyr Leu Cys His Phe Arg Thr Ala Leu Ile Leu
 5 10 15

Ala Val Cys Cys Gly Ser Ala Ser Ile Val Ser Leu Leu Leu Glu Gln
 20 25 30

Asn Ile Asp Val Ser Ser Gln Asp Leu Ser Gly Gln Thr Ala Arg Glu
 35 40 45

Tyr Ala Val Ser Ser Xaa His Asn Val
 50 55

<210> 566

<211> 55

<212> PRT

<213> Homo sapiens

<400> 566

Ile Leu Leu Glu Phe Phe Arg Asn Gln Arg Gly Ser Leu Asn Pro Arg
 5 10 15

Lys Thr Val Pro Phe Ile Lys Ser Glu Gly Gly Glu Lys Lys Gly His
 20 25 30

Cys Asn His Ser Val Val Ser Ile Asp Ser Ala Ala Ala Leu Leu Pro

000000"00000000

35

40

45

Leu Lys Leu Val Leu Leu Pro
 50 55

<210> 567

<211> 51

<212> PRT

<213> Homo sapiens

<400> 567

Tyr Ser Asp Phe Asp Val Phe Cys Ser His Thr Tyr Gly Tyr Met Leu
 5 10 15

Ser His Cys Ser Gln Ser Ser Ser Pro Leu Leu Trp Pro Leu Gly Ile
 20 25 30

Leu Thr Leu Ser Thr His Lys Met Ser Lys Leu Thr Leu Pro Pro Ile
 35 40 45

Phe Arg Thr
 50

<210> 568

<211> 75

<212> PRT

<213> Homo sapiens

<400> 568

Lys Val Gly Glu Tyr Ile Leu Gln Ser Leu Leu Arg Ile Arg Lys Ile
 5 10 15

Tyr Val Ala Phe Asn Ser Val Pro Ser Thr Cys Leu Leu Ala Ser Leu
 20 25 30

Thr Glu Thr Pro Val Thr Thr Ile Leu Thr Ile Ile Ile Asn Leu Thr
 35 40 45

Cys Phe Gln His Ala Glu Ser Ser Tyr Leu Phe Tyr Pro Leu Ala Asp
 50 55 60

Phe Leu Leu Gln His Ile Ser Leu Gly Lys Leu
 65 70 75

<210> 569

<211> 4809

<212> DNA

<213> Homo sapiens

<400> 569

005060" 6225960

gcatccagag tgggtggactg gttacaggct atgaacctac actgatgcgg caccaccacc 60
 cagagtccac rggttatggt ggttcacatt tactcttgct gtggtatggt ctataggttt 120
 ggacagatgt cggataatcc tttttacatt ttggcatcct tgggtagctc gtcttgtagg 180
 aatggacttg cttcaaagtg gaggcaggca gatccttcag acgggtatat ggagccctgt 240
 tttcagttgc ttttctaatt ctctcttata gtttacctca aaatcttcct gaggtctcgc 300
 ttctttttaa aatccttgct tactttgcag catcactctg acactcccat tgattcctca 360
 gcacctactg actacacggt taggagtgcagggttagaat tcatgtttta ttcattcttg 420
 ggtctgtagc acccagcaaa gtgctcagta aatgcgcagt aattgatttg acctctgaac 480
 aaatacacac tgtactaaga atctacacac cgaaagacaa aaacaagaca aatttgagtg 540
 ctacaggtgt cacgcttggt atcacacatg tgccctgtgta ttctcttagg tggttaccag 600
 gagctctgcc actgcatgtc cactagtgcagggttgcctc caccaccca gctgggtagc 660
 cgctgctctc acataagggt tccaattaaa attgccagga ataaattccc ccggactttg 720
 acttctcaag agctaagaag gtttgctgag tattctggca tgatgtttgg tgatcaaaca 780
 actgctggcc aaaaatgatg agtatttccc cctcttgctg aagatgtgct ccatacaata 840
 gtccatcaca ttcattctc atcagtctgg aagtgtgcag aacaacatgt aatagataat 900
 atgattggct gcacacttcc agactgatga atgatgaatg tgatggacta ttgtatggag 960
 cacatcttca gcaagagggg gaaataactca tcattttata tattacatgt tgttctgggt 1020
 tttttttttt tccaatgtcc agcctaaact ataaagtact ttgagaacgc acagtgaacc 1080
 ataagcttgc caataaagag tcctctgtgg tatggaactg gcttatttca tacacaatct 1140
 gcaacaatg agggcactat tggaaacata ctgtgctgca cagagcattt acaccgctta 1200
 tctttaatct tccccagcaa tccttgcttt gtgcgcattt atgatccttg ctctcagaag 1260
 tccacatact tttccccaac cgtaacaaat tatttaactc atctaatagt tgatgtccg 1320
 cgcagtctga aaacagtaat tgccttggg aagaagtgg ttaagagag ctctagggca 1380
 ctcatcacia ctccagccct gccctccatg tggtagcagc tctttggact ggggctaagt 1440
 gcttattctt gtgcttcatt cctggtaagc tcaatttctt taccttagga taactttgct 1500
 ggaaaagggc tcagattcag ccgaccattg tggcctctgt ggctgtcaca gcttgcctc 1560
 gacatgctat gatgttgggt ccccttctca tcccttggg atttctcttg ctggcccaca 1620
 gccagaacaa ctaggccttt tactccacca tccttttgtt ttcttttgtt tcgttggtaa 1680
 aaatcaatcc ttctaccatc catgcatagc aatttctaaa aactgaattt caagagcagt 1740
 atctgaagaa acaaacatga tttggctcct ttagtaaaaca gaataaattt taataaatca 1800
 actttgaaat agttgtaaga gttaagaaaa agcacaaaac tgagatcatc agagcagctt 1860
 ggctcctaaag gacaggcagc aggattctac agggtttgag ccttcctaag tgaagctgtt 1920
 tctgtaggc tcctgtctcc aagctcctag ctaacagccc cttctcccac gattggcaac 1980
 aaagagcaaa aataactttg tacttgatgc tgagtcagtg taaaaagcca taaaaattc 2040
 cctctaaatg tcaaaatgtt tgctctctt gaggtctctc tcctctact gggctctggat 2100
 aaattagcac tgggcttata ttgagtcaca gatctgggcc ctgccacaga gagcttctc 2160
 ctagtgtgtg atgctttttc tccaaactat tgatacaaaa tgcactggaa tagaaatcaa 2220
 cagaaaactg tcaaagggtg ggcatacaca ttctcatgta gatgtaaagc tgtgcttaga 2280
 attcctttgt ggagtctggt ttggtcttggt tttcttggt gtttgattca tttttttacg 2340
 taaattacaa aaacctcca catttcttca tggattgtat tagtccatgt tctccagaga 2400
 agcagaacga gttggatgta tgttttgga gagattatga ggaaccggt catgtgatga 2460
 aggaggttga gaggtcctgt gctctgccat ctgcaagctg aagacctgga aagctgaggg 2520
 tgtggtccca gtctgagctc gaaggcccaa gaaccagggg aaccaacggt gtagattcca 2580
 ggttgaaggc aggagaagat gtagtctcca gctcagcagg caggcaggaa gcaaatggg 2640
 taaattctc ctctctccac cttttgttcc attcaggcct tcaacagatt ggatgagcgc 2700
 cccccccc ccacactagg gagggccatc tgcttactg agtcggctga gtcaagtgc 2760
 agcctcatcc caaaacactc tccagacaca cgcagaaatg tttcatctgg gcaacctgtg 2820
 gccagtcatg ctgacacaca gaactaacca tgacatggat tcttcttaaa gcagtgatag 2880
 gagcgaacag aaacattttc ataattttca attattttta atgaaaacta tatctgatgg 2940
 aattgtttta acctagtctg gccacacatt atttctggg accgcccctc cttcaatccc 3000
 ttggacactg atgactttat gccagatta cactggaggc ctgtgctgat tttctaacac 3060
 atacctgcaa ctgagctggc aaaaagaaaa ctaggcaagt atgacagata catgatgcac 3120
 aggctaagt caaaggaaag aaaaacacca actgcaggga tgagggactc acccctttag 3180
 aagtttctac ttgagcagct agaagactac aatgccactc atcaaaacag tgactcaggg 3240

ggagtatttg ggataaagga ggaatctgat gttggaggtc aaatttgaag tgtctttaag 3300
 acctacaggt aacgagacag ctggacaaac acatggaact caggacaaag gctctaagga 3360
 cagcacagca gctgacatcc tgtgtgacag ccttgaaagc agcaggcccg ccgctcacat 3420
 tttggaaggg aaaatgggta caatgttgtc tgccactttg gggccttctt gggtcacatg 3480
 cattttacat ttatgcagtt gatatatatta tgtttcctgg gtcttttata cattagacac 3540
 catgattctc aatcctttgt tattttgtat taaaaaagc tgaattatta tttcaaatat 3600
 gggcaaatta gagccttcca tattgccaag gtgtatcaac cactgata ycaatgctc 3660
 tcttttgaat tagttttcca gttcacacct accatttatt tcatgattgg tttcagactt 3720
 gttcctcctg gaaacactcc ctaacaagca cccttgacag aatgaagaca caccacacac 3780
 atctacccca ttactgcatg tactcaagag tcagctttta tatgatctct cccaagtgc 3840
 cctataatgg ggatctttca ctcaccctaa agtgaggaca aaatacttga aagcatgagc 3900
 ccagtgcctg taggtgtgca attaacctca gaccaaggaa gtgccgaacg catctggctt 3960
 ttagcaaggc acctgacaaa gtcttcagg atgtttttgt acatgagcta gagaaatgta 4020
 cctggagaac agcttctact gccagatgat cttactcaaa agatgcagat taagcaaat 4080
 atcaacccaa aggggtgtcc ctgatggccc accagccct gtgcctggct cgtttcctat 4140
 gtttcctaga tttggtttca gacttgcctc tctgcagac actccctaac cagcatcctt 4200
 gcagaaaact ggtgaactag aaaaggcctg tgtgggtcac gtggccacc aacaccacag 4260
 cagtgtctaa ggtatgcgtg ggagcctgca cagcaggagc ggggtcttct ggagaccgc 4320
 atgagatgca aagggcagtg gacaaggagc caagggaggt ggctctagtc acgctggtat 4380
 ggtgccagct tgaggatgct gggcaagtcc cgagccgtct gccttcctag taccacagtt 4440
 accactgtct gttacctcgc gagtcaagt gcttcacgtg agacagctac gagacaggcc 4500
 cctggaaact ggaaaatgcy aagtaaagt catgcacaat tgttgttcac attttatctc 4560
 aatcactttt accaaatcag gctaaacct gggattcat aacgtcttg gctgtacaaa 4620
 ttgttccttg aaatgactca gagacatttt ctgaattggc ttccatcagc caagcatttc 4680
 ttcagaactg gaaaaatgct ttaaatttg ctttgtcatg attattaaaa cactctgtac 4740
 attttttatt attgaaatta acacattgcc tactttttta aaattggaaa aagaaaaaaa 4800
 aaaaaaaaa 4809

<210> 570

<211> 951

<212> DNA

<213> Homo sapiens

<400> 570

aaaattgaat attgagatac cattcttttag tgttaccttt tttaccaca tgtgtttctg 60
 aaaatatttg aattttattc atcttaaaaa ttggaccg ccttatttac catctttaat 120
 ccatttttagt actatgggtg agtacatgga attgaagtct ggcttaaatc ttcagaaagt 180
 tatatatcta ttttatttta tttttttgag acagagtctc gctgtgtcac ccaggctgga 240
 gtgcggtgcc acaatcttg ctcactgcaa cctctgagtc ccaggttcaa gcgatactca 300
 tgcctcggcc tcttgagtag ctgggactac aggcgtgcac caccacatct ggctaactct 360
 tttttgtatt ttttagtagag acggggtttc actgtggtct ccatctcctg acctcgtgat 420
 ccgctgcct cccaaagtgc tgggattaca ggcagagcc accgcacaca gctgggactg 480
 ggtaatttat aaagaaaaga ggtttaatga ctcacagttc cgcaggtctg gagaggcctc 540
 aggaaactta caatcatggt ggaaggcgaa ggggaagcaa ggcagctctt acatgggtggc 600
 aggagagaac gagtgagggg ggagactgcc aaaaactttt tttttttgag acaagagtct 660
 ggccctgttg cccaggtctg agtgagtggt catgatctca gctcactgca acctctgcct 720
 cacaggttca agcaattctc atgcctcagc ctcccgcata gctgggacca caggtatgca 780
 ccaccacacc tagctaattt ttgtagtttt agtagagatg gggctcact atgttgctca 840
 ggctggtcta aaactcctgg gctccagcaa tccgctgcc ttggcctccc aaagtgctgg 900
 ggttacaggc ataagccacc acatccagcc tgccacatac ttttaacta t 951

<210> 571

<211> 819

<212> DNA

<400> 571

<210> 572

<211> 203

<212> DNA

<213> Homo sapiens

<400> 572

tatagaataac	tcaagctatg	catcaagctt	ggtaccgagc	tcggatccac	tatttaacggc	60
cgccagtggtg	ctggaattcg	cccttagctc	ggatccacta	gtccagtggtg	gtggaattcc	120
attggtgttg	gccaacaca	atggagccac	cacatccagc	ctgccacata	cttttaaact	180
atcaggtctc	atgagaactc	atg				203

<210> 573

<211> 132

<212> PRT

<213> Homo sapiens

<400> 573

Met Val Glu Gly Glu Gly Glu Ala Arg His Val Leu His Gly Gly Arg

5 10 15

Arg Glu Arg Val Arg Gly Glu Thr Ala Thr Asn Phe Phe Phe Leu Arg
20 25 30

Gln Glu Ser Gly Pro Val Ala Gln Ala Gly Val Gln Trp His Asp Leu
35 40 45

Ser Ser Leu Gln Pro Leu Pro His Arg Phe Lys Gln Phe Ser Cys Leu
50 55 60

Ser Leu Pro His Ser Trp Asp His Arg Tyr Ala Pro Pro His Leu Ala
65 70 75 80

Asn Phe Cys Ser Phe Ser Arg Asp Gly Val Ser Leu Cys Cys Ser Gly
85 90 95

Trp Ser Lys Thr Pro Gly Leu Gln Gln Ser Ala Cys Leu Gly Leu Pro

```
<210> 574
<211> 62
<212> PRT
<213> Homo sapiens
```

```
<210> 575
<211> 76
<212> PRT
<213> Homo sapiens
```

```
<210> 576
<211> 68
<212> PRT
<213> Homo sapiens
```

<223> Xaa = Any Amino Acid

His Asp Ser Gln Ser Phe Val Ile Leu Tyr Tyr Lys Lys Leu Asn Tyr
20 25 30

<400> 581

Met Leu Glu Val Lys Phe Glu Val Ser Leu Arg Pro Thr Gly Asn Glu
 5 10 15

Thr Ala Gly Gln Thr His Gly Thr Gln Asp Lys Gly Ser Lys Asp Ser
 20 25 30

Thr Ala Ala Asp Ile Leu Cys Asp Ser Leu Glu Ser Ser Arg Pro Ala
 35 40 45

Ala His Ile Leu Glu Gly Lys Met Gly Thr Met Leu Ser Ala Thr Leu
 50 55 60

Gly Pro Ser Trp Val Thr Cys Ile Leu His Leu Cys Ser
 65 70 75

<210> 582

<211> 51

<212> PRT

<213> Homo sapiens

<400> 582

Met Leu Phe Leu Gln Thr Ile Asp Thr Lys Cys Thr Gly Ile Glu Ile
 5 10 15

Asn Arg Asn Trp Ser Lys Val Trp His Thr His Ser His Val Asp Val
 20 25 30

Lys Leu Cys Leu Glu Phe Leu Cys Gly Val Trp Phe Gly Leu Gly Phe
 35 40 45

Leu Gly Val
 50

<210> 583

<211> 60

<212> PRT

<213> Homo sapiens

<400> 583

Met Ser Thr Ser Asp Gly Phe Ala Pro Pro Pro Gln Leu Gly Ser Arg
 5 10 15

Cys Ser His Ile Arg Gly Pro Ile Lys Ile Ala Arg Asn Lys Phe Pro
 20 25 30

Arg Thr Leu Thr Ser Gln Glu Leu Arg Arg Phe Ala Glu Tyr Ser Gly
 35 40 45

Met Met Phe Gly Asp Gln Thr Thr Ala Gly Gln Lys
 50 55 60

000000"0225960

<213> Homo sapiens

<213> Homo sapiens

Leu Phe
50

<213> Homo sapiens

Ser Cys Arg Asn Gly Leu Ala Ser Lys Trp Arg Gln Ala Asp Pro Ser

35

40

45

Asp Gly Tyr Met Glu Pro Cys Phe Gln Leu Leu Phe
 50 55 60

<210> 587

<211> 1408

<212> DNA

<213> Homo sapiens

<400> 587

```

ctggacactt tgcgagggct tttgctggct gctgctgctg cccgtcatgc tactcatcgt 60
agcccgcccg gtgaagctcg ctgctttccc tacctcctta agtgactgcc aaacgcccac 120
cggctggaat tgctctgggt atgatgacag agaaaatgat ctcttctctt gtgacaccaa 180
cacctgtaaa tttgatgggg aatgtttaag aattggagac actgtgactt gcgtctgtca 240
gttcaagtgc aacaatgact atgtgcctgt gtgtggctcc aatggggaga gctaccagaa 300
tgagtgttac ctgcgacagg ctgcatgcaa acagcagagt gagatacttg tgggtgtcaga 360
aggatcatgt gccacagatg caggatcagg atctggagat ggagtccatg aaggctctgg 420
agaaactagt caaaaggaga catccacctg tgatatttgc cagtttggtg cagaatgtga 480
cgaagatgcc gaggatgtct ggtgtgtgtg taatattgac tgttctcaaa ccaacttcaa 540
tcccctctgc gcttctgatg ggaaatctta tgataatgca tgccaaatca aagaagcatc 600
gtgtcagaaa caggagaaaa ttgaagtcac gtctttgggt cgatgtcaag ataacacaac 660
tacaactact aagtctgaag atgggcatta tgcaagaaca gattatgcag agaatgctaa 720
caaattagaa gaaagtgcc aagaacacca cataccttgt ccggaacatt acaatggctt 780
ctgcatgcat ggggaagtgtg agcattctat caatatgcag gagccatctt gcagggtgta 840
tgctggttat actggacaac actgtgaaaa aaaggactac agtggtctat acgttgttcc 900
cggtcctgta cgatttcagt atgtcttaac cgcagctgtg attggaacaa ttcagattgc 960
tgtcatctgt gtgggtgggtcc tctgcatcac aaggaaatgc ccagaagca acagaattca 1020
cagacagaag caaaatacag ggcactacag ttcagacaat acaacaagag cgtccacgag 1080
gttaattctaa agggagcatg tttcacagtg gctggactac cgagagcttg gactacacaa 1140
tacagtatta tagacaaaag aataagacaa gagatctaca catgttgctt tgcatttgtg 1200
gtaatctaca ccaatgaaaa catgtactac agctatattt gattatgtat ggatatattt 1260
gaaatagtat acattgtctt gatgtttttt ctgtaatgta aataaactat ttatatcaca 1320
caatawagtt ttttctttcc catgtatttg ttatatataa taaataactca gtgatgagaa 1380
aaaaaaaaa aaaaaaaaaa rwmgaccc 1408

```

<210> 588

<211> 81

<212> PRT

<213> Homo sapiens

<400> 588

Met Pro Gln Lys Gln Gln Asn Ser Gln Thr Glu Ala Lys Tyr Arg Ala
 5 10 15

Leu Gln Phe Arg Gln Tyr Asn Lys Ser Val His Glu Val Asn Leu Lys
 20 25 30

Gly Ala Cys Phe Thr Val Ala Gly Leu Pro Arg Ala Trp Thr Thr Gln
 35 40 45

Tyr Ser Ile Ile Asp Lys Arg Ile Arg Gln Glu Ile Tyr Thr Cys Cys
 50 55 60

009060" 6225960

Ile

<211> 157

<213> Homo sapiens

Met Thr Met Cys Leu Cys Val Ala Pro Met Gly Arg Ala Thr Arg Met
5 10 15

Trp Cys Gln Lys Asp His Val Pro Gln Met Gln Asp Gln Asp Leu Glu
35 40 45

Met Glu Ser Met Lys Ala Leu Glu Lys Leu Val Lys Arg Arg His Pro
50 55 60

Pro	Val	Ile	Phe	Ala	Ser	Leu	Val	Gln	Asn	Val	Thr	Lys	Met	Pro	Arg
65					70					75					80

Met Ser Gly Val Cys Val Ile Leu Thr Val Leu Lys Pro Thr Ser Ile
85 90 95

Pro Ser Ala Leu Leu Met Gly Asn Leu Met Ile Met His Ala Lys Ser
100 105 110

Lys Lys His Arg Val Arg Asn Arg Arg Lys Leu Lys Ser Cys Leu Trp
115 120 125

Val	Asp	Val	Lys	Ile	Thr	Gln	Leu	Gln	Leu	Leu	Ser	Leu	Lys	Met	Gly
130						135					140				

Ile Met Gln Glu Gln Ile Met Gln Arg Met Leu Thr Asn
145 150 155

<211> 347

<213> Homo sapiens

Met Leu Leu Ile Val Ala Arg Pro Val Lys Leu Ala Ala Phe Pro Thr
5 10 15

Ser	Leu	Ser	Asp	Cys	Gln	Thr	Pro	Thr	Gly	Trp	Asn	Cys	Ser	Gly	Tyr
			20					25					30		
Asp	Asp	Arg	Glu	Asn	Asp	Leu	Phe	Leu	Cys	Asp	Thr	Asn	Thr	Cys	Lys
		35					40					45			
Phe	Asp	Gly	Glu	Cys	Leu	Arg	Ile	Gly	Asp	Thr	Val	Thr	Cys	Val	Cys
	50					55					60				
Gln	Phe	Lys	Cys	Asn	Asn	Asp	Tyr	Val	Pro	Val	Cys	Gly	Ser	Asn	Gly
	65				70					75					80
Glu	Ser	Tyr	Gln	Asn	Glu	Cys	Tyr	Leu	Arg	Gln	Ala	Ala	Cys	Lys	Gln
				85					90					95	
Gln	Ser	Glu	Ile	Leu	Val	Val	Ser	Glu	Gly	Ser	Cys	Ala	Thr	Asp	Ala
			100					105					110		
Gly	Ser	Gly	Ser	Gly	Asp	Gly	Val	His	Glu	Gly	Ser	Gly	Glu	Thr	Ser
		115					120					125			
Gln	Lys	Glu	Thr	Ser	Thr	Cys	Asp	Ile	Cys	Gln	Phe	Gly	Ala	Glu	Cys
	130					135					140				
Asp	Glu	Asp	Ala	Glu	Asp	Val	Trp	Cys	Val	Cys	Asn	Ile	Asp	Cys	Ser
145					150					155					160
Gln	Thr	Asn	Phe	Asn	Pro	Leu	Cys	Ala	Ser	Asp	Gly	Lys	Ser	Tyr	Asp
			165						170					175	
Asn	Ala	Cys	Gln	Ile	Lys	Glu	Ala	Ser	Cys	Gln	Lys	Gln	Glu	Lys	Ile
			180					185					190		
Glu	Val	Met	Ser	Leu	Gly	Arg	Cys	Gln	Asp	Asn	Thr	Thr	Thr	Thr	Thr
		195					200					205			
Lys	Ser	Glu	Asp	Gly	His	Tyr	Ala	Arg	Thr	Asp	Tyr	Ala	Glu	Asn	Ala
	210					215					220				
Asn	Lys	Leu	Glu	Glu	Ser	Ala	Arg	Glu	His	His	Ile	Pro	Cys	Pro	Glu
225					230					235					240
His	Tyr	Asn	Gly	Phe	Cys	Met	His	Gly	Lys	Cys	Glu	His	Ser	Ile	Asn
			245						250					255	
Met	Gln	Glu	Pro	Ser	Cys	Arg	Cys	Asp	Ala	Gly	Tyr	Thr	Gly	Gln	His
			260					265					270		
Cys	Glu	Lys	Lys	Asp	Tyr	Ser	Val	Leu	Tyr	Val	Val	Pro	Gly	Pro	Val
		275					280					285			
Arg	Phe	Gln	Tyr	Val	Leu	Ile	Ala	Ala	Val	Ile	Gly	Thr	Ile	Gln	Ile
	290					295					300				

Ala Val Ile Cys Val Val Val Leu Cys Ile Thr Arg Lys Cys Pro Arg
305 310 315 320

Ser Asn Arg Ile His Arg Gln Lys Gln Asn Thr Gly His Tyr Ser Ser
325 330 335

Asp Asn Thr Thr Arg Ala Ser Thr Arg Leu Ile
340 345

<210> 591

<211> 565

<212> DNA

<213> Homo sapien

<400> 591

actaaagcaa	atgaacaagc	tgacttgcta	gtatcatctg	cattcattga	agcacaagaa	60
cttcatgcct	tgactcatgt	aaatgcaata	ggattaaaaa	ataaatttga	tatcacatgg	120
aaacagacaa	aaaatattgt	acaacattgc	accagtgctc	agattctaca	cctggccact	180
caggaagcaa	gagttaatcc	cagaggctca	tgtcctaata	tggtatggca	aatggatgct	240
atgcacgtac	cttcatttgg	aaaattgtca	tttgtccatg	tgacagttga	tacttattca	300
catttcatat	gggcaacctg	ccagacagga	gaaagtactt	cccatgttaa	aagacattta	360
ttatcttgtt	ttcctgtcat	gggagttcca	gaaaaagtta	aaacagacaa	tggggccaggt	420
tactgtagta	aagcatttca	aaaattctta	aatcagtggg	aaattacaca	tacaatagga	480
attctctata	attcccaagg	acaggccata	attgaaggaa	ctaatagaac	actcaaagct	540
caattgggta	aacaaaaaaa	aaaaa				565

<210> 592

<211> 188

<212> PRT

<213> Homo sapien

<400> 592

Thr	Lys	Ala	Asn	Glu	Gln	Ala	Asp	Leu	Leu	Val	Ser	Ser	Ala	Phe	Ile
1			5					10					15		
Glu	Ala	Gln	Glu	Leu	His	Ala	Leu	Thr	His	Val	Asn	Ala	Ile	Gly	Leu
		20						25					30		
Lys	Asn	Lys	Phe	Asp	Ile	Thr	Trp	Lys	Gln	Thr	Lys	Asn	Ile	Val	Gln
	35						40					45			
His	Cys	Thr	Gln	Cys	Gln	Ile	Leu	His	Leu	Ala	Thr	Gln	Glu	Ala	Arg
	50					55					60				
Val	Asn	Pro	Arg	Gly	Leu	Cys	Pro	Asn	Val	Leu	Trp	Gln	Met	Asp	Val
65					70					75					80
Met	His	Val	Pro	Ser	Phe	Gly	Lys	Leu	Ser	Phe	Val	His	Val	Thr	Val
			85						90					95	
Asp	Thr	Tyr	Ser	His	Phe	Ile	Trp	Ala	Thr	Cys	Gln	Thr	Gly	Glu	Ser
			100					105					110		
Thr	Ser	His	Val	Lys	Arg	His	Leu	Leu	Ser	Cys	Phe	Pro	Val	Met	Gly
		115					120					125			
Val	Pro	Glu	Lys	Val	Lys	Thr	Asp	Asn	Gly	Pro	Gly	Tyr	Cys	Ser	Lys
	130					135					140				
Ala	Phe	Gln	Lys	Phe	Leu	Asn	Gln	Trp	Lys	Ile	Thr	His	Thr	Ile	Gly
145					150					155					160

00529.00600

```
<210> 593
<211> 271
<212> DNA
<213> Homo sapien
```

[illegible]

```
<220>
<221> misc_feature
<222> (1)...(376)
<223> n = A,T,C or G
```

```
<210> 595
<211> 242
<212> DNA
<213> Homo sapien
```

```

<400> 595
agncgtgctgn tcgtnccectn tatgtggcct catnntgagg acaanagtng cactgaggct      60
tqngnatgcc aggcaaggnc aagctggctc aaaaagcatc caccacctc tgnaanggggt      120

```

atgccangag cangtgcacc agtcccaact angagncccn ggcatgntac atctttcttcc 180
 acccctnaaa ntttngngcta caangnccat ttttcttttt ctcttaaggg ncncttggt 240
 tc 242

<210> 596
 <211> 535
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(535)
 <223> n = A,T,C or G

<400> 596
 accagttaga tactgctaaa nagatattta tgcagcctca tatgttaagt cgtatatttt 60
 gaaagctttt taaatttttt ctttaagaag atttttagatg cttatcactg agtaccagag 120
 ggatgtaggc tgatgccctt atcaacaaag tcagggactg tggcacacaa ggattgacta 180
 ctgcagacac ggccacaatg ctacctctag agggcctgaa tccccctgcc ctctctgggtg 240
 gggagaaggg ctggcagagc cattagcatg ggctccggcc aatcctggcc actttgacac 300
 tcctgggtgct gaccaggggt cctggaggaa gggatgaggt gggcagtaga gatgctcagg 360
 gcagtggccc ctttccatcc acactggaac tatttcagta ttttaccacc aattcagcca 420
 ttcccttggtg cgctggctga acatcagccc tgctccaggt ctcagtttcc cctttgtaaa 480
 gggaaagctc tggattcagg gagtgatgaa gaggtcatca tgggtcttgag aattc 535

<210> 597
 <211> 257
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(257)
 <223> n = A,T,C or G

<400> 597
 tttcnatacc caaaantacc ccatattang accanacatt tgtctnggaa aaattaccat 60
 tntntaant ttgggccacc tgagannaaa tgggtgtaat ncatgataag atggancagn 120
 attnctctta agatnngatn agaccccggt tttcaccgaa catatccaag naccacaatag 180
 gnaacaagcc acgggnggag tcacaaacat atattcttta ctctcataat ccgtnnacaa 240
 naactnttgn acttgac 257

<210> 598
 <211> 222
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(222)
 <223> n = A,T,C or G

<400> 598
 nntggntacc gtcnaaactt nnttggttac ccgagctcgg atccactagt ccagtgtggt 60

```
<210> 599
<211> 238
<212> DNA
<213> Homo sapien
```

<400>	599						
gcatgcacatc	ancgatgtnt	ttggnnacct	ganattngct	aaaactngng	natgccgggn		60
atgnaggttt	ggtantgatc	tatgcactga	catactatgg	ggagctttca	tgtggagtgn		120
tcgacaangt	tgctnгансn	gagaagtga	gatctcagtt	gaaagggtca	tgtgaataca		180
cnttacacgt	gaaaaagaag	cacattggga	atatcacgaa	acgnccacca	acatcctg		238

```
<220>
<221> misc_feature
<222> (1)...(232)
<223> n = A,T,C or G
```

```
<210> 601
<211> 547
<212> DNA
<213> Homo sapien
```

<400>	601						
catttgtgttg	gggaaaaaat	gatttgtata	agcagtgggg	ctatttgcca	ttgctttttt		60
tttttcttaa	atatcaccta	ttaggttgaa	aacctgaaat	tgcagctttc	tgtagaaatg		120
gcggaagaca	aactaacatt	tttaaagcgc	tctcatttag	ctctgatgag	tactacaccc		180
ctnatattct	tctgatacta	aaataatttt	cctagtgtag	tctaaacttt	tttaaaaaga		240
catgtaatcc	gcggagttag	taactcaaaa	cgagtgcata	tnggaagtat	cgcagccggt		300
nctggatnaa	attcccagct	tgcctngett	ctnagccggg	gggcggtnaa	aaaaacatct		360
gcagcccngg	ggnaaaaacc	ttcgacattgt	tcttacgtgt	ttacgttatt	ttatttcctt		420

nnagcaaggc nggganttgg ggactcgaaa tggtagagtt gggctgggga tgcacctgtg 480
 tacataaaaag ncgtccagaa gagggacggt tacaggcngg ganctccaaa ggtcagtcctc 540
 tgccatt 547

<210> 602
 <211> 826
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(826)
 <223> n = A,T,C or G

<400> 602
 cgggggggnt tacgtctctc tggacgcttt tattgtacca gggcgatccc agcccaactg 60
 taccattcga gtcctactc ctgccttgct ctagggaat aaaataacgt aaacacgtaa 120
 gaacaatgcg aaagcgtttt cttccctagg ctgcagattg tcttcttcac cgcccctgct 180
 tagctagcta gctagctggg aatttaatcc agaaacggct tgcgatacct cctagatgca 240
 ctcgttttga gttacaaact ccgcggatta catgtctttt taaaaaagtt tagactacac 300
 tagggaaaat tatttttagta tcagaagaat atcagggggg gtagtactca tcagagctna 360
 atgagagcgc tttaaaaatg ttagtttgtc ttccgccatt tctacagaaa gctgcaattt 420
 caggtttttca ncctaatagg tgatatntaa gaaaaaaaaa acaatcgcan atagcccact 480
 gctttttacaa atcattttttc tcttctaggt atagcctgtc aggtggccta atgtattttt 540
 gacatctcta ggaattttta tagaccagaa atgggtgcca gagatatgcc tgcactaatc 600
 ttaagtgggg atttatgtat ttctcaanca agtgattaaa gcaaaactag gcacgaatga 660
 aatcaagatc tttaggccag aaatcatgaa nanttttana attattttan gaatctgtgg 720
 cttctcttct taaaatngaa aaaaaaattg tttaaaccca naaggtctga ataccaagc 780
 nccctgaacn anagaacaan gccggagcac cccctcccaa atcccc 826

<210> 603
 <211> 817
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(817)
 <223> n = A,T,C or G

<400> 603
 nnangacttt tgtggtntta tacaattntt ttttctatth ctatgaagag aaagccacag 60
 agtctaaaaa taattctaaa actcatcatg actttcttgc ctaaaagatc ttgatttcaa 120
 tegtgcctag ttttgcttta atcacttgct tgagaaatac ataaatcccc acttaagatt 180
 agtgcaggca tatctctggc acccatttct ggttctatta aaattcctag agatgtcaaa 240
 aattacatta ggccacctga caggctatac ctagaagaga aaaaatgatt tgtaaaaagca 300
 gtggggctat ttgcgattgc tttttttttt tcttaaatat cacctattag gttgaaaacc 360
 tgaaattgca gctttctgta gaaatggcgg aagacaaact aacattttta aagcgtctc 420
 atttagctct gatgagtact acaccctga tattcttctg atactaaaat aattttccta 480
 gtgtagtcta aactttttta aaaagacatg taatccgcgg agtttgtaac tcaaaacgag 540
 tgcacttagg aggtatcgca agccgtttct ggattaaatt cccagctagc ttgcttgctt 600
 agcaggggag ggnaaanaag acatctgcag cctagggaag aaaacctttc gcattgttct 660
 tacgtgttta cgttatttta tttcctanaa caaggcngaa ttgggactcg aatggttcag 720
 ttgggggtggg ggatccccctg gtncataaaa ngtcanaaag anggtacagg cggaacncca 780

817

```
<220>  
<221> misc_feature  
<222> (1)...(694)  
<223> n = A,T,C or G
```

```
<210> 605
<211> 678
<212> DNA
<213> Homo sapien
```

<400> 605							
taaaaatcta	gactacacta	ggaaattatt	ttantatcag	agaatatca	ggggtgtagt		60
actcatcana	gctaaatgag	agcgctttta	aaatgttagt	ttgtcttccg	ccatttctac		120
agaaagctgc	aatttcaggt	tttcaacctt	ataggtgata	tttaagaaaa	aaaaaaagca		180
atcgcaaata	gccccactgc	ttttacaaat	catttttttct	cttctaggtt	tagcctgtca		240
ggtggcctaa	tgtaattttt	gacatctcta	ggaattttta	tagaaccaga	aatgggtgcc		300
agagatatgc	ctgcactaat	cttaagtggg	gatttatgta	tttctcaagc	aagtgattaa		360
agcaaaaacta	ggcacgattg	aaatcaanat	cttttaggca	agaaagtcat	gatgagtttt		420
anaattatttt	taggactctg	tggttttctc	ttcatagaaa	tagaaaaaaa	aaattgtata		480
aaaaccacaa	aaggtcctga	atagcccaaa	gcaacactga	acaaaangaa	caagcagga		540
agcaacacac	taccggaatt	caattatact	accaaggtgt	antaacaaa	acagcattct		600
attgggcata	aaatgagacca	aagaccagtg	ggaaacagaa	taaagaancc	caaaataaat		660
cctatatatta	cnqcccn						678

```
<210> 606
<211> 263
<212> DNA
<213> Homo sapien
```



```

<220>
<221> misc_feature
<222> (1)...(263)
<223> n = A,T,C or G

<400> 606
gtgggggtcng cancagccaa ctcagcttcc tttcgggctt tgtttagcaga cggatcatcc      60
tctagtccac tgtgntcaaa ttccattgtg tggggggcnc tcgcctcggc canagatctg      120
agtgancana cntgtcccca ctgaggtgcc ccacagcngn ttgtnttcag canggggctna      180
caactcgacc ggcagcgnaa ggctggcaga antgngcgcc tnnctcattc ctacgcngtn      240
ngccgcagga aggangacag gcc                                263

<210> 607
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 607
ccatgtgggt cccggttgtc tt                                22

<210> 608
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 608
gataggggtg ctcagggggt gg                                22

<210> 609
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 609
gctggacagg gggcaaaagc tggggcagtg aacctgtgc          40

<210> 610
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

```

<400> 610
ccttgtccag atagcccagt agctgac 27

<210> 611
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 611
gatagagaaa accgtccagg ccagtattgt gggaggctgg gagtgc 46

<210> 612
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 612
gcacatgggt cactgccccca gcttttgccc cctgtccagc 40

<210> 613
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 613
gccgctcgag ttagaattcg gggttggcca cgatggtg 38

<210> 614
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 614
cggcgggcat atgcatcacc atcaccatca catcataaac ggcgaggact gca 53

<210> 615
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 615
gcactcccag cctcccacaa tactggcctg gacgggttttc tctatc

46

<210> 616
<211> 1350
<212> DNA
<213> Homo sapien

<400> 616
atgcatcacc atcaccatca catcataaac ggcgaggact gcagcccgc ctcgcagccc 60
tggcaggcgg cactggtcat ggaaaacgaa ttgttctgct cgggcgtcct ggtgcatccg 120
cagtgggtgc tgtcagccgc aactgtttc cagaactcct acaccatcgg gctgggcctg 180
cacagtcttg aggccgacca agagccaggg agccagatgg tggaggccag cctctccgta 240
cggcaccag agtacaacag acccttgctc gctaacgacc tcatgctcat caagttggac 300
gaatccgtgt ccgagtctga caccatccgg agcatcagca ttgcttcgca gtgccctacc 360
gcggggaact cttgcctcgt ttctggctgg ggtctgctgg cgaacggcag aatgcctacc 420
gtgctgcagt gcgtgaacgt gtcgggtggtg tctgaggagg tctgcagtaa gctctatgac 480
ccgctgtacc accccagcat gttctgcgcc ggcggaggggc aagaccagaa ggactcctgc 540
aacggtgact ctggggggcc cctgatctgc aacgggtact tgcagggcct tgtgtctttc 600
ggaaaagccc cgtgtggcca agttggcgtg ccagggtgtct acaccaacct ctgcaaattc 660
actgagtgga tagagaaaac cgtccaggcc agtattgtgg gaggtctggga gtgcgagaag 720
cattcccacc cctggcaggt gcttgtggcc tctcgtggca gggcagctctg cggcgggtgtt 780
ctggtgcacc cccagtgggt cctcacagct gccactgca tcaggaacaa aagcgtgatc 840
ttgctgggtc ggcacagcct gtttcacct gaagacacag gccaggtatt tcaggtcagc 900
cacagcttcc cacaccgct ctacgatatg agcctcctga agaatcgatt cctcaggcca 960
ggtgatgact ccagccacga cctcatgctg ctccgcctgt cagagcctgc cgagctcacg 1020
gatgctgtga aggtcatgga cctgcccacc caggagccag cactggggac cacctgctac 1080
gcctcaggct ggggcagcat tgaaccagag gagttcttga ccccaaagaa acttcagtgt 1140
gtggacctcc atgttatttc caatgacgtg tgtgcgcaag ttccacctca gaaggtgacc 1200
aagttcatgc tgtgtgctgg acgctggaca gggggcaaaa gctggggcag tgaaccatgt 1260
gcctgcccgc aaaggccttc cctgtacacc aaggtggtgc attaccggaa gtggatcaag 1320
gacaccatcg tggccaaccc cgaattctaa 1350

<210> 617
<211> 449
<212> PRT
<213> Homo sapien

<400> 617
Met His His His His His Ile Ile Asn Gly Glu Asp Cys Ser Pro
1 5 10 15
His Ser Gln Pro Trp Gln Ala Ala Leu Val Met Glu Asn Glu Leu Phe
20 25 30
Cys Ser Gly Val Leu Val His Pro Gln Trp Val Leu Ser Ala Ala His
35 40 45
Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu Gly Leu His Ser Leu Glu
50 55 60
Ala Asp Gln Glu Pro Gly Ser Gln Met Val Glu Ala Ser Leu Ser Val
65 70 75 80
Arg His Pro Glu Tyr Asn Arg Pro Leu Leu Ala Asn Asp Leu Met Leu
85 90 95
Ile Lys Leu Asp Glu Ser Val Ser Glu Ser Asp Thr Ile Arg Ser Ile
100 105 110

```
<210> 618
<211> 385
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(385)
<223> n = A,T,C or G
```

```

<400> 618
ctgtgctgag aaccaaagc tatgancact gcttttccaa atgtccataa naccaacatt      60
tttatcacta ccaccatcac ctgggagctc nttagaaagc tagtctcccg ggcaccaccc      120
tggcctactg aacctaattgt gcattttaaca agattnacgt ngaaatctgc aaagcacagg      180
ggcngataac agtaccacct gntctgggtc ctanccccan gacccttaca gtctaactgg      240
gacacaaggg cttnaaatca aattgcctat cattaagata tacaanganc ntgagaaact      300
gctncaactta tntattaagg ngctctaaga cttagaaacn aaangcantg ctgagangat      360
tcaaatatga ngggggnac tttnc                                           385

```

<210> 619

<211> 869

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(869)

<223> n = A,T,C or G

```

<400> 619
gatatcccg ggaattcgcg cgcgctcgac ctctacttgt ttagacataa atgcagtcta      60
gcattaaaga tcctttaaaa aaatgttttc ccaatgggta aaagacaagc tcaaataaat      120
gaactctcat acatatgcca aaattgatga gtagataaat atttcagtag gtagttacta      180
gctttctgtg tatgagtaaa catatgggag aaatttaaaa cactaaagta gactcaatga      240
aagcatagta tcctatgtat tcgtttttca gaaatgtcta atgaaggag gaaacaatga      300
atgaatgccc ttattcctct tagagtgtcg ggacatgggt ttgcctgaaa acttcatgtg      360
aattttatat tttgctacac attacacca tcttagactt atacgtataa gacataaggc      420
atatcttatg tcttacatgt ataataatct aagcagaaca aaaaataacg aaatattttc      480
ttccccaat ttttgagaca gatggatttt ccggaaaagat gtgttttagct tttaatcctg      540
tggttttgtg taccacctgg cacactagag tgttgctcta attcagttag ttgtaactct      600
gggtgaacag tggaaatact aggggtacatt ttaaaaaatgc taatgctcgg gctcgcgtga      660
agaccaaatt aattggaatc tctgngggng gnattgatct ttttataatc tttctanang      720
attctaattg gcttccaggg atgaaaaccn ctgntggagc tnggaacctt cctttagttt      780
ggagaaaccc cgatgagggt ntnttaggcn ccgcctnttt ttggcctggg cttccccctt      840
tatntntttt tggaanggnc cnaattttt                                           869

```

<210> 620

<211> 339

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(339)

<223> n = A,T,C or G

```

<400> 620
gngcgggctt cncggtgctt gctctcgctg ccgacgctct ttttccacca gctgtaggan      60
aagcccgaag accactgggc ccccggttag cccaagtacc actggctctc ctggctcctg      120
acgctnccgg tcttctcgtt ggcgtagact gccagcttcg gagaccctc agcccctccc      180
cgcttttctc caccocagga ggccatcagt agcgagctac tgcctcggcc acaacctccc      240
agcangatag ccccggtttt ccaatctgcg aaaggaggac cgccnagccc gaaatgccna      300
gccagcnat cactgccacg ccgagccnag cgctcgtgc                                           339

```

```
<220>
<221> misc_feature
<222> (1)...(681)
<223> n = A,T,C or G
```

<400> 623
 aaaactgtac tcgcgcgctg catgtcgaca ctagtggatc caaagaatcg gcacgagcga 60
 aaangctcan gcagcccggc tggcgcgcgc cgctcctccc cccaggaaag ccaangtggg 120
 ngctgatgtg gctgcangag ctogtttcac agccccctcan gtgganctgg ttgggcccgcg 180
 gctgccangg gcggaagtgg gtgtccccc cangtcagccc caaggctgcc cctcaciaaag 240
 cactgggtgg ttgcctccac tgcacacctg ggctccgaac ccgctcccct gctgtggang 300
 cccaccgtgg gaatccaggt ccccagggtg actgcctgcc ttgcctcac tgcacctct 360
 gcccacactt ccctgcctag anaccgggaa ggggtgtgt cggtantggg gccacactgg 420
 atgtggcagc accgactgtg ggggtggacc tggccttgcc ggggtgcaaaa gtggggggccc 480
 ngggaaaagc acctgaagtg gccctgaaaa atccccctt aatttttccc caatttgggg 540
 ctcaacaaa aggaaattgc tgaagccaan ggtaccaagg tcacccctaa ggccagggtg 600
 aaaaggtccc aaaattccaa tccccacct ttgggcttnc ctcttggaac cccggccccc 660
 tctcntgaan ttttaaaaaa n 681

<210> 624
 <211> 661
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(661)
 <223> n = A,T,C or G

<400> 624
 attggtctta ctgtaccacc ggggtggaaat cgatggccgc ggcgtctaaa tatccgattt 60
 tttttttttt tcctcttctg actgtccatg gacaaatgaa actaacttaa totaactaaa 120
 aaacacaact atatttttgaa gattttctat ctgcactcaa ggacactttc caacnccgtt 180
 ttgttacctt ttggtcttgt ctctgaacat gaaattnatc tcaagggtt ngattttctgg 240
 acctcctatt cctgctatgg gtttgatatt tcttgggctc cagggccact gttgcattgg 300
 gntgacagnt acctcctagc ccatanctc ctatcttggg aaacaaacct aacaactacg 360
 tgtaccttcc atagatctct gattgagtct cagtatnccg ttgctcatgg gcgattcact 420
 tgaatccgtt attggtgcca acaatcctga ctcatgggnn aatggatcct atcacgttcc 480
 cctgattngc aacccttgta tacatanatc taatcgcata gaatctagcn tnggntatgc 540
 gcggctacgc tatcagggtg tgntaactat ngcatggcta cgaancctga tcatgatcna 600
 gggctcatgga ctcttatcag ggggggttggg ccgngcttct ttttcnnacc ttggtaaaac 660
 c 661

<210> 625
 <211> 181
 <212> DNA
 <213> Homo sapien

<400> 625
 gcaacaatca gatcatgtta aagtaaatct ccattgccct ggatcacttc aggattttaat 60
 tgtccaagga gagcagggtt ctctgtgtaa aaaaagggtg ggaaatgttt gagagtaaaa 120
 aatacaaaaat tcaaccggtc gaaaatacac cactccattc agtgctctac ccccataagc 180
 c 181

<210> 626
 <211> 181
 <212> DNA
 <213> Homo sapien

```
<210> 627
<211> 813
<212> DNA
<213> Homo sapien
```

<400>	627						
accaagctgg	agctcgcgcg	cctgcaggtc	gacactagtg	gatccaaagt	gaacgtgaag		60
gtgagcagag	gagaacttgc	gatggcaaag	ttaaaaacaa	gaggagatga	tggtcttggt		120
gtggcacagg	atgttaaaaa	aattctcctg	tccttaagga	gttactgcta	tttgagtaat		180
gtgccacttc	cctacatagc	cttctatgca	gaaatgctat	atttccactt	cacaaccag		240
aacgtgcatt	ttatttttaca	tttagaggag	gaacaaacaa	ccagaaggca	aaaactgggtg		300
cattatTTTT	tgcaattctc	ttggaaagag	ttcgttttta	acttctgctc	agacagcaca		360
caactactgg	gaatatatTT	taattttcaa	tctgatgtgt	gacatctggg	aactcattta		420
ttgctaata	agtttttcaca	ggaagcagca	gtcaccagta	gctcatctta	tttttcagtt		480
ggcaaagtgt	tgttttacctt	ttattggcct	gcattcgggtg	ctcttatcac	aggatatTTA		540
attagaaaac	gcaagtagcc	taacatagaa	nagaaatgga	gtggtagata	atagtagata		600
gaatggctaa	atattttttat	tcacgtgatg	taatatcact	gnaatttatg	gttaaaaaatt		660
atgtaatact	caaaaggaat	tctcagactg	gcgaaacagc	tggncaacag	ctntcacagg		720
gctttnanct	cctnttgagc	tttccccctg	ntggacttta	gtcttctctt	tacncccgna		780
gttnccattn	nttaccaatt	gtncggggaa	ana				813

```
<220>
<221> misc_feature
<222> (1)...(646)
<223> n = A,T,C or G
```

<400>	628								
tttgggnggn	ggtgtctcnt	ttgggtggac	tttttgggtc	gtagggcccc	aaggccgtta				60
atcccgtaat	aacggaagac	gaagaagagt	cagaagagtg	cttctataag	gatcgggacg				120
agactacctt	agaggaataa	aggaaaaaag	cagaggagga	agagtggtag	aaggagtcag				180
aagaaaccca	cacgtcgttc	tgaacctgga	gccttatcaa	aaaggtctag	ataaacgata				240
gcgatctcga	tatcgagctc	aagaggtagg	tttagagact	tctcgtcctc	gagagcgaaa				300
tggaagatct	cgacgacgat	aagaagttaa	agtgtagagg	gtgcttgagg	agcgcgtgga				360
aggattctgc	ggaggggacc	atcgacgtag	agacttgaag	gcctactaag	gtccacaaga				420
agccccgcta	tttctccgaa	tggtcggagc	gtacagtagt	cgacgtcgat	cggcagacaa				480
gctggcggtc	gactcgaagt	gttcgggcga	atcgacttat	aatagtcgcg	cgctagtaac				540
qtaggaacac	gaagagtagt	cgaaagaaaa	cgtttagtga	gggaaaagat	tagggaaaaa				600

<221> misc_feature

<222> (1)...(526)

<223> n = A,T,C or G

<400> 631

```
ccntcggtt ggggtttttt ctgagccccc cccccccccc cccccccccc cccccccggc    60
cccatagccc caccggnccc acccaaatTT taacaaaata aatntaccta tcgntcacct    120
atcccnogta tcngtaggt cggtaaccgg accgngatc ncnacgattn ttcgggtcgt    180
cncccttaan acggncccg agccnccgga anaaatacta cgagngactc taatntagca    240
anaccgcgcg tcnattanta gcacccttag tcttccaatg ncnnggattn ngaatccttn    300
naagttatcg ggtagaacgg gtcccggtcc cccgccctct ttncaattaa cgcgggtac    360
aaantcggtt tctaaattcc ncacgaattt ngncggcaac attcnccggg ccttattanc    420
cntttccaac cccgatacnc nagctcgatc gggctttanc gaatccgggg tcncccccca    480
ngantccggg tccctttgagt ngctctagga cggttacgac ggagga                    526
```

<210> 632

<211> 647

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(647)

<223> n = A,T,C or G

<400> 632

```
tttggngggc gggngctcat ttgggtggac tttttgggtc gtaggaacct ggtatgaggg    60
gtgttttgag tttcttcttc gtcgtctctg ggaggttcgg ttccgattga gattcggggt    120
cgtctttatc ttacgaggca ccctgatatt gttgcgcttt ggtttggttg tggagagttt    180
tgtctactc tagcgggtca tgcggatgat atgtagcctg cgtggcctga tagtgatgtt    240
gtgagcttga gaggggagtt gtgggtgttg cgggcggagt aggaggggtt ggagcaccgg    300
gattgggaga tatagaatca taagtgttag gtataggctg attgagcgag ttcgtggaat    360
tcgtgtggtc atcataatta gagtgaggat gggctctata tttcttagag gacgcacggg    420
cgtgattcgg ggtttgatgg gtgttcttct tgtgggcacg attagcttgt tcatgatggg    480
aaggaccata ctgtttcgaa tgaggattcg tgtcttcgga ttgttggtga tattgtggnc    540
tanactatTT agtgtaagcc ggaggtgggt tgccgtgggt gagtatccga nnttcattcg    600
ganggtatgc gtgcggagcg gtcctttaga acattccgga aaaatgg                    647
```

<210> 633

<211> 630

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(630)

<223> n = A,T,C or G

<400> 633

```
tccttcgggt tgggtttttt tctgaccccc cccccccccc cccctcggga aggcctctag    60
gtcccccacc gtctctctaa tctcaggaa ccgatccacc caaccaactt actaatgtcc    120
tacagtaaac acccgagaat ataaacccac acctaggcct ccaatcctac caggaagca    180
agaagccgta gtctagcgta ttacgaaccc gagatagaga cggagatact tagttttatt    240
ctctcgggaat aggaagacg actggggagg gaatataggc tagcgcgggg ataggggcta    300
```

```
<210> 634
<211> 647
<212> DNA
<213> Homo sapien
```

<400> 634						
ccntcggcctt	gggtttttttt	ctgacccccc	ccccccccc	cctccactaa	gancttaacc	60
caaccctata	gtttactcgt	ataggggaat	cgaggagaaa	taggaacgaa	gagcgggtga	120
taaagagaaa	gtactttcct	ttatatgtta	agagcttagc	gtaatgactt	tcgttatatg	180
gctagttagt	tttatccggc	gttatagggc	ttagttctgg	ttatctcggg	tctaattccc	240
ttagtatgct	cgggagttta	acgaggtcac	gggatagcgc	gtaccctttc	taaggttcct	300
ggaaagctat	tcgttattta	tcgcgattct	cgaggtcgaa	aggatcaagg	atcttccctt	360
ttactaccct	agtcgggtta	gcggtcggtc	aaaactagt	tagtaacctt	acctcctoga	420
aagttatagc	cgaaacaacg	tattagtcga	aattatagcg	gatagatcga	gacgggttctt	480
tctcgggttc	tcagcccgga	atccctctat	ttgggggtct	tctccctctt	cccctttgtc	540
ttccgcctta	gcttccaagg	ttcctcgga	gcgaggggtt	ctacttaagt	cgntagecgtt	600
ccttataaac	cncctacagg	cagacccctt	tgtaaacggc	tcgggggt		647

```
<210> 635
<211> 645
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(645)
<223> n = A,T,C or G
```

<400>	635						
ccttcggtt	gggtttttt	ctgagcccc	cccccccc	cccgaaactc	gccttaccct		60
agatacccaa	agaatagttc	cactcaactt	cgtctaagta	aaactctaga	acttccaaac		120
ataaaaagact	tgcgcggtt	agctacacag	cctacgggaa	tctcacgaat	cccgattcaa		180
gtcccactct	cgaccacacc	ccggtatcgt	cgttttccca	taccaatgtc	gaaaaataaa		240
ataaaaatcca	gtcaagcccc	acggtaaagc	ggggtagggc	taggcgaaga	ggcaggaacc		300
gttcgaggcc	gggggctttc	aaaatacaaa	acaactactt	aaagtttacc	ccttctaagg		360
tcgggggcaa	cggttaaagc	acgcctctaa	agtactactc	gtttcgagaa	ggggtagtca		420
tctcccgcat	agagactctc	gcgtatatca	actcgcacgc	cttctagcat	tccgacggtc		480
gcccgcggct	acatatcttg	cggattagct	ccgagggact	atagggttaa	ttagtctagt		540
aaattctctt	agaggatagt	gggggtcgta	gttaggcagt	acgaggggac	atggnctgcg		600
tcgtgctcta	ccttgacagc	atactcttat	aaacatcttt	ttcct			645

<210> 636

```
<220>
<221> misc_feature
<222> (1)...(606)
<223> n = A,T,C or G
```

```
<210> 639
<211> 592
<212> DNA
<213> Homo sapien
```

<400>	639								
tccttcgggt	tgggtttttt	tctgagcccc	ccccccccc	cccccgga	cgagaaaaca				60
atcccaccct	accgcgggga	gtgggttgna	cgcttagttc	tagaatcctc	ggaatcgctc				120
tccggcggtt	gtagtccgg	cgattccgag	tatgccgaag	tgtatcgctc	cgtctagagg				180
ttggtatctg	tttatcgcg	tgacgctatt	gactcggatg	ctttcgaagt	agggggatag				240
gcgcatagat	acgcctccgc	ggtgtcctct	gaagtggccg	catccgtgga	cgcagcgtag				300
acagctctgg	tggacgataa	cggtctctcg	tactcctact	cggctatta	tgttagagag				360
gacttgtttc	tgaacggata	taccattagc	gaaggggtac	ctccgcataa	cgcaggcgtt				420
tctaacagtt	cttcggggcg	ctccgaattc	agattgacgc	ctccgcagca	ttgtgggatc				480
ctcttcggtt	agccctcttt	ataggatttc	tcctccgcc	cgaaganggg	ctggtcgtcc				540
ccggcangta	tgtctagctc	gaacgctttg	ttactccttt	gttttcgaaa	na				592

```
<220>
<221> misc_feature
<222> (1)...(637)
<223> n = A,T,C or G
```

<400> 640							
ctttgtggcg	gtggntgtct	catttggggtg	gacttttttgg	gtcgtaggct	tatccgggtn		60
gggctcccg	agtagcttag	gatcgccggc	tagttccggt	cccgcccgtc	gaaagcgcg		120
ttcggcgggc	ggcccccgct	tcgttcgcg	gctttaccct	catagagtgc	caggtctcgg		180
ttcttacggg	ttcgtcggcg	atagatttta	cggcgagagg	tcggtatctt	cgccgcttta		240
cgttcggtcg	gcatctacgc	ctagttcaca	ggtagtttat	gcgccggagc	gcgtgacgga		300
gaggttatac	gggacgcgga	agaaccgcct	ccaaatgact	agtacaggct	cgttcgggcg		360
taatatctct	cqctcggtcg	gcggttctta	cttctagggc	cgctctacgg	tttaaggcgg		420

```

tcgttagatc ttagaaacta tactcaagtt tcagtcggaa gaaaggaagt agagagaagg 480
gtaaacgatt acctccggtt ctagcccttt ttactcgcat aacgggagaa cggggtccgg 540
ctctcagata cgcctcgca gacgtcgca ttcaacttta acctccgcta gggcatccgt 600
atacgggttaa cgcggtaaaa gcgacctcg aaacctc 637

```

```

<210> 641
<211> 649
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(649)
<223> n = A,T,C or G

```

```

<400> 641
ctntgtggcg gtggttgtct cagtttgggt ggatttttgg gtcgtaggna acctggtatg 60
aggtctagtt tcttcaacga ttcttggttc agttacgca cctatcctt atcttacaat 120
gtcttctaca tcaggttcat caattaatat atcaattaca cattaacgac ggtgtgacgc 180
aatatgagaa agtatacatt aagggtatta tatattattc gcttaaaaag gttcctgaca 240
tgggacaact tcaccacca ttctagaagc cccccctcct gtaggacccc ctcgagttcc 300
ccattatctt agttcagttt tcattttttaa accaggaggg tatcggtttt taataggtac 360
tattttgtca aacttttcag aagctttatc ttcaaataata cttgcaccat ctgtactagg 420
agcactaact attcgagtct attacagctc aacagaaaat aattgaaatt aaacaaccta 480
agtatcgctc accataaccc catcgggctc tcaccccat tcttcataag ttctagagca 540
tcctgagctc tttcctatta cccttgatgg tactcatggg ctaatacccc ccgcagttat 600
aggtccttat ggatcctatg ctaccaccgg tctaatacct tctatcacn 649

```

```

<210> 642
<211> 645
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(645)
<223> n = A,T,C or G

```

```

<400> 642
tccttcgggt tgggtttttt ttcgctcgcg gttactatta tcgattgtta cttgtaaagg 60
cgatactccc accgctcacg atattagacc tgctcctcta gaagcgaacg gcgataggtc 120
tactcggcgg gcgaagacgg cgaacgggta ggaggagcca tatgcaaccc taacggagat 180
tataagtact gggaaaaata ctagtattaa ggtagcgggt taagataggt ggagagacac 240
tattcacgag cataagcact tagaaggtct tctcgaggag aggtaggcta cggactacgt 300
tccttcttcc tctagcctcg agagggagta tagatgattc gcaaaaagaga atccctccta 360
taogctggca taactagacg acgcgtcgtc gggaaatctc gccaaccccta ttgcgacctc 420
caaaaggaag attgtcgttt catagaacgc taatactccg ggtcttcccg aatcatagcc 480
gcataatcggt aagaagacgg taaaatcgcg cgattctaac aagattctgt agacttaagg 540
ctaagcacta gaagcgatct cgattccgga tcttaagatc atactaatag ttcggtcaca 600
ccagacgacg attagccact agaagcccta ctccgtngaa accgg 645

```

```

<210> 643
<211> 586
<212> DNA

```



```
<210> 646
<211> 645
<212> DNA
<213> Homo sapien
```

<400>	646									
tcttcgcggt	tgggtttttt	tctgagcccc	ccccccccc	ccccacgcc	aagtacacag					60
accaccaaaa	aacaacgtca	acacaacttc	gggtatacgg	accttaagag	agacccccta					120
gtagacccta	ccacagccat	ccaatagtca	aacaacaagg	gcgcacccaa	tccatccata					180
gagctatcaa	acaacgaggg	ggaaaggaaa	gagcaggggt	aacttagcag	agatcggaagt					240
cggcactaat	tcttttcaag	tactcgctcg	gcttgtagtt	cggggtaaag	tccgctctca					300
aagggccaac	gaggttttaa	agcgaccccc	gtatcgagtc	ttcttcgtat	tcattaagggc					360
gttaaaggta	cgagacctag	aagagagtag	aattagccca	ccaaatcgcc	taaacgggca					420
aaaacgacca	aaagtcaaa	accettacaa	atatcacctt	aaaacgcxaa	ccccaaaaac					480
gcgatcagta	acgcacgtac	ctttcccacy	cttttctttc	tttcactctc	caaaaacaaac					540
ccgaatatatt	agcgcaaaaa	atatccgagg	gagaattaga	agctattacc	cgaaaaaaaaa					600
antgaanangq	ntaaatnqt	qqqqaatana	cgtttggttt	ttctg						645

```
<220>
<221> misc_feature
<222> (1)...(753)
<223> n = A,T,C or G
```

[illegible]


```

ctggactccc cgcggtggcg gccgctctag aactagtgga tccgtgttgg ctcaattctc      60
aaggetgttg ctgtgcggcc tgttccccac acgtgctgct cagctcaggc aagcaccgag      120
cttgtgttgt ttcattgctca gcgtggaggc cctcctcca ggtcgtgct ctgtgggggt      180
cccatacact caggctccta ggaggagtcc atttagaaag ccagggtttt tctcagagtc      240
ttagttcctt gtgctgtcat ccatttcaca cgacttgggc cctgctcggg gcaacacagc      300
aagagaaaag acagggaaaa taagagaggg accttgcaca cacacgctct ggaccacaga      360
gccctgtgcc cagctcctct gtcaatacag gtggaatctc gtgcaggatc gcaggggtct      420
gtgatgccac caaagagcag gccgggacag ggtaggaga gaaaggagag ggaagtgggg      480
gtttctccta cgcactctta ttgacagagg gaaaggcggg tttgtattgg ggtgtcgggt      540
ctttgcaccc acngcacagt tgtgagacac cccatcctn agatcaaagc cccacatata      600
gcttggggaa aaacaaaacn aaacaaaaca aaaacagtaa acctccatgc canttgttgg      660
gnaagttttn aatttncttc ccnaccan cttgcttc      698

```

<210> 659

<211> 750

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(750)

<223> n = A,T,C or G

<400> 659

```

ncaantcggg ctcaccgcg gtggcgccg ctctagacta gtggatctc ctcatgggcc      60
tgatattctc tgaacatat atgaacattg cttatgaaaa attatttgta ngaaaattgt      120
gaggcctaag aatgntattt tcttttagtg atggtctttg ttgtctctg taaggnaactt      180
gtgggcactc gtaagcttgg atctctttaa tctaatacca gntttgagat tttcttggcc      240
ccatagatga attaaaactg gcgtacttct tgtttacaag anggataagt ctctagggt      300
aagtcttttg gggctccaag tcaaaaagat gagggattta ccagttctct aaccttggta      360
gccccagact ccaaactttg cttctagtc ccaagaggct atcaaaaagc aaaggccatc      420
ttccaccttc ttttccanaa cagcacacat tccagacagt acttgaaagc aggaacctcc      480
ttatccctta aaaacctctt ggaancatct tccctctctt gcttctacta tgcttggccc      540
acctancatt cncntttttc tggaaaccgg aaaaancttn tgacttnngt tggctacatt      600
cagcttggcc cctacaatn tggtttccat ctgccctaan gaaattttta agggcacttt      660
ttttntggcc cctgactttc nntttttagg gctttcccc angctttgcc cctttggtta      720
aaggggttat tttcttccc cttttggaag      750

```

<210> 660

<211> 849

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(849)

<223> n = A,T,C or G

<400> 660

```

tcggatccac tagtccagtg tgggtggaatt cgcgccccgc gtcgacgggc agtagtggtg      60
tgcntntcta aatgttataa ttatttcaga attactctgc cagaaagtta tgatcatata      120
tagaagagtt tgtagctaac tttgaaagta gtggaaagtg gttttcatgt attgtttggg      180
ttaatttaat ttgtattata tttgggtttt agttcaggta atttttttgt tgaaaacttc      240
aatgacaat ttcttcattg ttactaaaga tcactcatgt ggagtagttt cagatttttt      300

```


<220>
 <221> misc_feature
 <222> (1)...(694)
 <223> n = A,T,C or G

<400> 665
 cttttcaaat cattttttnct cttctaggta tancctgtca ggtggcctaa tgtaattttt 60
 gacatctcta ngaattttta tagaaccaga aatgggtgcc agagatatgc ctgcactaat 120
 cttaagtggg gatattatgta tttctcaagc aagtgattaa agcaaaacta ggacagattg 180
 aaatcaagat ctttttaggca anaaagtcac gatgagtttt agaattattt taggactctg 240
 tggctttctc ttcatagaaa tagaaaaaaa aattgtataa aaccacaaaa ggtcctgaat 300
 agccaaagca acactganca aaaagaacan agcaggggaag caacacacta ccngaattca 360
 aattatacta ccagggtgta gtaacacaaa cagcatttcta ttggcataaa atagacacca 420
 agaccaatgg ancagaataa agaacccac aaataaatcc atatatntac cgccanctga 480
 ttatcaataa cnaacaccaa gaacatatnt taagggacnt nctattcaat aantagtgtc 540
 ggnaaaaact gggaaatcca tatgcagaaa naatgaaact agacccttat ccctcaccat 600
 acgcaaannt caacttcgga atgggattac aaaacttaag acattccaac ccaagaaact 660
 atnaaancta ctattaagaa aacagatcnc nccc 694

<210> 666
 <211> 705
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(705)
 <223> n = A,T,C or G

<400> 666
 tttaaaaatt tagatacact angaaaatta ttttagtattc agaagaatat caggggggtgt 60
 agtactcatc agagctaaat gagagcgctt taaaaatgtt agtttgtctt ccgccatttc 120
 tacagaaagc tgcaatttca ggttttcaac ctaatagggtg atatttaaga aaaaaaaaaa 180
 gcaatcgcaa atagcccccac tgcttttaca aatcattttt tctcttctag gtatagcctg 240
 tcaggtggcc taatgtaatt tttgacatct ctagggaattt taatagaacc agaaatgggt 300
 gccagagata tgcctgcact aatcttaagt ggggatttat gtattttctca agcaagtgtat 360
 taaagcaaaa ctaggcacga ttgaaatcaa gatcttttag gcaagaaaagt catgatgagt 420
 tttanaatta ttttaggact ctgtggcttt ctcttcatag aaatagaaaa aaaaattgta 480
 taaaaccaca aaaggtcctg aatagcccaa gcaacactga acaaaaagaa caaagcagga 540
 agcaacacac taccagaatt caaattatac taccaggtg tagtaaccaa aacagcattc 600
 tattgggcnt aaaatagacc naagaccaat ggaacagaat aaagaaccca aaataaatcc 660
 atatttttac agccagctna ttatcaataa aaacnccaag aacnt 705

<210> 667
 <211> 817
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(817)
 <223> n = A,T,C or G


```

<400> 667
nnangacttt tgtggnttta tacaattntt ttttctattt ctatgaagag aaagccacag      60
agtcctaaaa taattctaaa actcatcatg actttcttgc ctaaaagatc ttgatttcaa      120
tcgtgcctag ttttgcttta atcacttgct tgagaaatac ataaatcccc acttaagatt      180
agtgcaggca tatctctggc acccatttct ggttctatta aaattcctag agatgtcaaa      240
aattacatta ggccacctga caggctatac ctagaagaga aaaaatgatt tgtaaaagca      300
gtggggctat ttgcgattgc tttttttttt tcttaaatat cacctattag gttgaaaacc      360
tgaaattgca gctttctgta gaaatggcgg aagacaaact aacattttta aagcgtctct      420
atttagctct gatgagtact acacccctga tattcttctg atactaaaat aattttccta      480
gtgtagtcta aactttttta aaaagacatg taatccgcgg agtttgtaac tcaaaacgag      540
tgcacttagg aggtatcgca agcogtttct ggattaaatt ccagctagc ttgcttgctt      600
agcaggggcg ggnaaanaag acatctgcag cctagggaag aaaacctttc gcattgttct      660
tacgtgttta cgttatttta tttcctanaa caaggcngaa ttgggactcg aatgggttcag      720
ttgggggtggg ggatccccctg gtncataaaa ngtcanaaag anggtacagg cggaacncca      780
agggtcgctcc tgcatttana ctcggaattt tggtgcc                                817

```

<210> 668

<211> 826

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(826)

<223> n = A,T,C or G

```

<400> 668
cgggggggnt tacgtctctc tggacgcttt tattgtacca gggcgatccc agcccaactg      60
taccattcga gtccctactc ctgccttgct ctagggaat aaaataacgt aaacacgtaa      120
gaacaatgcg aaagcgtttt ctccctagg ctgcagattg tcttcttcac cgccccctgct      180
tagctagcta gctagctggg aatttaatcc agaaacggct tgcgatacct cctagatgca      240
ctcgttttga gttacaaact ccgcggatta catgtctttt taaaaaagtt tagactacac      300
tagggaaaat tatttttagta tcagaagaat atcagggggg gtagtactca tcagagctna      360
atgagagcgc tttaaaaatg ttagtttgct ttccgccatt tctacagaaa gctgcaattt      420
caggttttca ncctaatagg tgatatntaa gaaaaaaaaa acaatcgcan atagcccact      480
gctttttaca atcatttttc tcttctaggt atagcctgtc aggtggccta atgtattttt      540
gacatctcta ggaattttta tagaccagaa atgggtgcca gagatatgcc tgcaactaat      600
ttaagtgggg atttatgtat ttctcaanca agtgattaaa gcaaaactag gcacgaatga      660
aatcaagatc tttaggccag aaatcatgaa nanttttana attattttan gaatctgtgg      720
cttctcttct taaaatngaa aaaaaaattg tttaaaccca naaggtctga ataccaagc      780
nccctgaacn anagaacaan gccggagcac ccctcccaa atcccc                                826

```

<210> 669

<211> 547

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(547)

<223> n = A,T,C or G

```

<400> 669
cattgtgttg gggaaaaaat gatttgtata agcagtgggg ctatttgca ttgctttttt      60

```


<223> n = A, T, C or G

actagtcacg	tgtggtggaa	tccattgtg	ttgggagcag	tttaaaaaaa	aaaaagacna	60
aatatacnac	tcttgatnaa	acataaaggt	acagtgggtct	atgaggaana	gaaaaggtac	120
ctnaggatgc	aaaantacct	accacatggg	aaccgttngt	ccacactcat	tccnnanaaa	180
accgagtcct	ctcanttnca	cacgtgtacg	tttcagttgg	gaagtgcctg	ccattactcc	240
naagcctaga	accttcacgt	cctgaagggt	ctggaagggt	tttcagattg	cttaaganac	300
gnggccttc	catattcttc	tccactaccc	nggggaacgg	aacaaatgga	gctgcgacng	360
ggaagcgctc	cttcccntcc	gaacgccttc	tttcaaacct	gcctgccttc	cnggcgaatg	420
gaccggaagg	tttncntngct	tcttttcanc	ccnaattact	tctctgngttg	aaattttggcc	480
tgttggtttg	caaatcgngg	aatttgttta	cttctntcat	gtcctgtgtgt	ggnncnaaccg	540
gctcncctgt	tgccctccctt	tngaaagggt	ttcatcaggc	ccgcgccctt	ctcttntaan	600
ngtcctaate	cggncnggac	cactcgggga	aaattttttc	ttttcgaaaa	gccgccccnt	660
cgtccggct						670

<213> Homo sapien

<223> n = A, T, C or G

actagtccag	tgtggtggaa	ttccattgtg	ttgggagtag	gtctactaca	nctacttcc	60
cctatcatan	aagancttan	caacnttcat	gatccccccc	tctannccct	tttctcanc	120
tgcttcttag	tctgtttgt	cctnttctta	acantcntaa	ganagatnac	taatnctact	180
atctctnacc	tccggaanct	acaanacgtc	tggaactatt	cngaccccat	gcancncat	240
ntcccatcgt	cctcccagcc	cctncccttc	ctttacntta	ctnaacgaag	gtcgacgatc	300
cctcccntac	ctcccnnncc	attgggnccc	aangnactg	gacctcacga	ntacaccnac	360
tacggggnga	ctaagnctgn	aactccttac	atatntcccc	gttaccceccn	gaacncagcg	420
aacnqcnaca	ccttggacnt	caagaanta				449

<213> Homo sapien

<223> n = A, T, C or G

tttcngtgtg	gtggaattcg	cggccgcgtc	gacgagaaga	nggaggagga	naaggagaag	60
gagaagaagg	agaanaagga	ggagaaggag	aagaaggaga	agaaatcatc	atcatcatca	120
tccactgtct	ngcaactatt	taagtttgcn	antcccttga	aaacaggtac	ttttgtttca	180
atgtttggga	ccactnctga	cnatgannag	aanaccaata	aatgcttgat	naatgaaaaa	240
nccacttttt	acctgttaga	accctgaggc	taagagaant	gatgtgactc	gacttagtta	300
ccacaaacta	tgatcctaqc	atnaattggg	gcattctaac	acctcaactc	cctgtgcaag	360

attacacatg tttactacaa gagatgttna taagtaaaga aggcctgata tacaatctaa 480
cagacnantg agataaatct taantcacia ctgacntccc ttttggggcg g 531

<210> 686
<211> 336
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(336)
<223> n = A,T,C or G

<400> 686
ggngncctna tgagcgcgcg taatacgatc atatagggcg aattgggtac cgggcccccc 60
tcaagaacac tacaagctat gtcctcttct canagagccc tgaantttta acatattgaa 120
agctctnalc ttgccaaana actccactta acttcaaaac acaccctcca cacacatcat 180
gatcaactna gatcttactg aaccagaatc ctnaatggca tacttcagga acaggggtcc 240
anagaagcag ttctcaaant gcagctnaaa aagaaactga aaaccaatt catgcaanac 300
ctagggttta tttgagagca ttttccagtg cagatt 336

<210> 687
<211> 271
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(271)
<223> n = A,T,C or G

<400> 687
aatctgcact ggaaaatgct ctaaaataag ccctagggtct tgcattgaatt ggggttttcag 60
tttcttttta agctgcactt tgagaactgc ttctctggac cctgttccct gaagtatgcc 120
atthaggtt ctgggttcagt aagatctcag ttaatcatga tgtgtgtgga ggggtgtgtt 180
tgaagttnag tggagtctt tggcaagatc agagctttca atatgttnaa acttcagggc 240
tctctgagaa gaggacatag cttgtagtgt t 271

<210> 688
<211> 740
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(740)
<223> n = A,T,C or G

<400> 688
tgatgaagcg cgcgtnttac nactcactat nggggcgaan tatgggtacc gggnccccct 60
cgaagcggcc gccctttttt tntttttttg tgagagttaa aataaaatat ttgagttaa 120
tttaaagttt gagtttaatt aaaatatatg gcatatccca agttgggctt tgcanaaaga 180
acacttctca ggaactgtta gttggtgtac caggaactca gaagggtcct gttattaaat 240
atatttgga aatgcatgga ttctctgaan atcncctctgc atgtgagcaa cacttacatc 300

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(882)

<223> n = A,T,C or G

<400> 691

ttactcacta	tagggctcga	gcgggccgctg	aattctgctg	cagtgagctg	tgattatgtc	60
cctgcactcc	agcctggatg	acagaacacg	atcatttctc	taaagacaaa	caaaaaacat	120
aaaataaaaac	tagtataagg	atagaagccc	agggttgatt	taagtctgcg	gaaatcataa	180
accataggtc	agacttctca	ttgatgaggt	acttggtggg	tagaatacaa	ttaggatat	240
ttggtctaga	aaccaggatg	gaattagaga	ataaaaagact	gagcaatagc	atgttatagt	300
attagaaata	ctatagaaat	aggaaaagcc	ctgattatga	ctttggagtt	ctgatccaac	360
atctgggatt	atthagatat	tttaaaggaa	aacgatgact	tttagctctc	aggatgttag	420
tttctcctcaac	cataaaatga	agagcctcga	aaagatttgc	tttaccagat	tatttctgaa	480
gtcaattcca	gttctaaaat	tccatcactg	ngcactaagg	caaattgaat	tgaataaagt	540
attgggnatg	cataaaatac	tctattttta	aaaangaata	gtaattatcc	attggnaaca	600
gacgcantca	tccagncatc	tcctaccctg	ncccatgnnc	tatgtagana	tgtanctcta	660
atcccttaac	aaaccgattt	tgcaaaggag	cttanccttg	gggtacttgg	tcanggcaac	720
tggtctactt	tnaagactca	tcttctactta	ctgggcacca	aatnccctacc	attgcatcaa	780
actgggggttc	ccatncaagg	caaaccctgn	gaaatcttta	atcccgaat	tggcgcctcaa	840
ttttgngggg	tttccnaaaa	gaatcntccc	ccccgagggg	cc		882

<210> 692

<211> 235

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(235)

<223> n = A,T,C or G

<400> 692

ccgcactngt	aangnccgcc	agngngctgn	aantccgctn	agcncggatc	cactagtcca	60
ttgatggtaa	aagggtagct	tactggnatg	tccgnctgct	ccanganata	atacncagga	120
cttctcanag	cacttaatat	gttaatataa	aactncngna	aaaaagatnt	tcnatgaanc	180
nttcctctta	ggaggtcagg	ngagaatagt	gttaatgnca	ttaagganag	aacga	235

<210> 693

<211> 383

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(383)

<223> n = A,T,C or G

<400> 693

nttatgtaag	aaatgtcata	tatctttttat	tttcttttaaa	tcaaaaataaa	tatgactttg	60
agcatcccat	cccatgcccc	atcctatcag	aatggtagga	acatcaacac	aaataattag	120
taatgcaccg	catctacatt	cccattgctct	ctttacttct	tcagcattgc	ctaaaggcat	180

caacctacaa gctctctaat catgctcacc taaaagattc ccgggatcta ataggctcaa 1680
 agaaacttct tctagaaata taaaagagaa aattggatta tgcaaaaatt cattattaat 1740
 ttttttcatc catcctttaa ttcagcaaac atttatctgt tgttgacttt atgcagtatg 1800
 gccttttaag gattggggga caggtgaaga acgggggtgcc agaatgcac ctcctactaa 1860
 tgaggtcagt acacatttgc attttaaaat gccctgtcca gctgggcatg gtggatcatg 1920
 cctgtaatct caacattgga aggccaaggc aggaggattg cttcagccca ggagttcaag 1980
 accagcctgg gcaacataga aagaccccat ctctcaatca atcaatcaat gccctgtctt 2040
 tgaaaaataa actctttaag aaaggtttaa tgggcagggt gtggtagctc atgcctataa 2100
 tacagcactt tgggaggctg aggcaggagg atcactttag ccagaagtt caagaccagc 2160
 ctgggcaaca agtgacacct catctcaatt ttttaataaa atgaatacat acataaggaa 2220
 agataaaaag aaaagttaa tgaaagaata cagtataaaa caaatctctt ggacctaaaa 2280
 gtatttttgt tcaagccaaa tattgtgaat cacctctctg tgttgaggat acagaatctc 2340
 taagcccagg aaactgagca gaaagtcat gtactaacta atcaaccga ggcaaggcaa 2400
 aaatgagact aactaatcaa tccgaggcaa ggggcaaatt agacggaacc tgactctggt 2460
 ctattaagcg acaactttcc ctctgttgta ttttctttt attcaatgta aaaggataaa 2520
 aactctctaa aactaaaaac aatgtttgtc aggagttaca aacctgacc aactaattat 2580
 ggggaatcat aaaatatgac tgtatgagat cttgatgggt taaaaagtgt accactgtt 2640
 aatcacttta aacattaatg aacttaaaaa tgaatttacg gagattggaa tgtttctttc 2700
 ctgttgattt agttggctca ggctgccata acaaaatacc acagactggg aggcttaagt 2760
 aacagaaatt catctctcac agttctgggg gctggaagtc cacgatcaag gtgcaggaaa 2820
 ggcaggcttc attctgaggc ccctctcttg gctcacatgt ggccaccctc ccactgctg 2880
 ctcacatgac ctctttgtgc tcttgaaaag aggggtgtggg ggacagaggg aaagagaagg 2940
 agagggaaact ctctggtgtc tctctttca aggaccctaa cctgggccac tttggcccag 3000
 gcactgtggg gtgggggggt gtggctgtc tgctctgagt ggccaagata aagcaacaga 3060
 aaaatgtcca aagctgtgca gcaaagacaa gccaccgaac agggatctgc tcatcagtgt 3120
 ggggacctcc aagtcggcca ccctggaggc aagcccccac agagcccatg caaggtggca 3180
 gcagcagaag aagggaattg tccctgtcct tggcacattc ctcaccgacc tgggtgatgt 3240
 ggacactgcy atgaatggt atgtggatga gaatatgat gactcccaga aaaggagacc 3300
 cagctgtctc ggtggctgca aatcattaca gccttcaccc tggggaggaa ctgggggccc 3360
 ggttctgggt cagagagcag ccagtgagg gtgagagcta cagcctgtcc tgccagctgg 3420
 atccccagtc ccggtcaacc agtaatcaag gctgagcaga tcaggcttcc cggagctggt 3480
 cttgggaagc cagccctggg gtgagttggc tctgtctgtg gtactgagac aatattgtca 3540
 taaattcaat gcgcccttgt atcccttttt cttttttatc tgtctacatc tataatcact 3600
 atgcatacta gtctttgtta gtgtttctat tcmacttaat agagatatgt tataacttaa 3660
 aaaaaaaaaa aaaa 3674

<210> 699

<211> 2051

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(2051)

<223> n = A,T,C or G

<400> 699

ggaccagggg ctgaagtga cccccagcac agcacagctg ctctataaaa acgtggccag 60
 actttttttt ttgaagcaag tccctgttct tgttcgtcct gactagtccc atcagggcc 120
 tggatcccaa gactcagcat ccaaggctcc ctccaggaat cctggcagct cagcactatt 180
 tatcctgttt catctgagag caaaaatgta aaattggatg cacagaaaag tgactcaaag 240
 tgcttaatat ctagaagaaa tctaggagca gcaagaagag caggacaaac aggccaggcg 300
 gtgtcaggag ccaggtctc cagctggang gaacgtcaac cctgcagtgg gagcaggggc 360
 cctttgcaca tctaggcac agatggtaat gtagacacca caggtaagct gggcttggt 420


```

cctgtgtctc tagccagccc accactggag ctatcaagcc agtagcaact cagcagttcc 900
ttggacagag cttccaggag caaatgaaat cctttctgcc actgcctttg cagtgaactg 960
cccttgctat cctcagaaga tatatcacgg gagcaaagac cctaagtgcc atatcaacac 1020
ctccaataag ctgcagttga cccaaagaac aagccaatcc atctcccaca ggttccacac 1080
acactccact actcatcacc agacagggaa ccttggettg ggcccacagc acagaccctc 1140
catcctgggc cgattacact gagtgattgc taactcacat gtctctggga tggagcacc 1200
aggagacaag caaagtgggt gagcagcaag tcaggatgat tggagcccag agggcaggga 1260
gagctatctc tctgggctcc acttgccctt gtgagacact ttgtcccagc actccttagt 1320
ctgcttgcc tctccagggc cccagcctgg ccacacctgc ttacagggca ctctcagatg 1380
cccataccat agtttctgtg ctagtggacc gtaccatata agtggagagc tgcagcaagg 1440
tggcccntac ggccacgcac cagcctgcac attacctctc catactgcag ccctttatat 1500
ggaaacttcc tacatcactt tgctgtgtgt gtttacacag gtggattttg ctttacttgc 1560
actgacagca cacaggaggg cagcacacac cccaaccac atcaactgcc attaaagaaa 1620
agaaatttca gcccataatt tcatgtccag caaatttagg catcataagt gaaggagaaa 1680
taagatcctt ttcagacaag caaatgctga ggaattcaa tatcaccaga tctaccttac 1740
aagagctcct gaaggaagca ctaaatatgg aaagaaaaaa ccatcaccag cactacaaa 1800
aatgcagtga agaacgcagt gaattacgca gtccagtgat gctaaaaacc aaccacatac 1860
gttaagtctg caaaataacc agctgacagc atgacgacag gataaatcca cacataccat 1920
tactaacctt aaatgaaaat gggctaaatg ctcccattga aagacatggg gcaagctgga 1980
taaagaacca agaccactg gagtatgctg tcttcaagaa acccatctca catgcggtgg 2040
catacatagg ctcaaaataa aggaatggag aaaaatattt caagcaaatg gaaaacagaa 2100
aaaagcaggt gttgcactcc tactttctga caaacagac tatgcgaata aagataaaaa 2160
agagaaggac attacaaagg tggctctgac ctttgatata tctcattgct tgataccaac 2220
ctgggctggt ttaattgccc aaanccaata ggataatttg ctgaggttgt ggagcttctc 2280
ccctgcagag agtccttgat ctcccaaat ttggttgaga tgtaagggtg attttgcgtg 2340
acaactcctt ttctgaagtt ttactcattt ccaaaaagga aggcaagttt tctgcttcc 2400
atgacgatgg agagcaggca tctcctttcc tgagtttcag cttgcttctg acaggggaagg 2460
tgagtgtgag ttttttccag cttctaagat ggcagagaaac gatcaccagc ctgagcctta 2520
tttccaggta agtagctgaa ttagagtttt gtcttaaaat ttttctttaa tgattaaaat 2580
gtaagattac ccaccagctg cttttaattt ctcccttagc attagaacac tcagtaatca 2640
tatgaattgt gcatttggtt gttttgctta actctttctg tttgtttatg tttggggttt 2700
tattgttggt gtttcacttt tctcccatct ctctctgact tgggtcaaat caaaggaatg 2760
ttcgaaattg tggggagcaa ggcattctgaa atggctaaaa ctctgtggc tgcaaaaaat 2820
agaaataaaa aaaaaaaaaa a 2841

```

<210> 701

<211> 3228

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(3228)

<223> n = A,T,C or G

<400> 701

```

tccgccccat tgacgcaaat ggcggtaggc gtgtacggtg ggaggtctat ataagcagag 60
ctctcnggct aactagagaa cccactgctt actggcttat cgaaattaat acgactcact 120
ataggagagac ccaagctggc tagcgtttta acttaagctt ggtaccgagc tcggatccac 180
tagtccagtg tgggtggaatt ccattgtgtt gggcaggaaa caagcaaagt ggtggagcag 240
caagtcagggt gatgtggagc ccagagggtca gggatggctg tctctctagg gtccacttgc 300
ccttgtgaga cactttatcc cagcacttta ggaatactga ggtcatacca gccacatctt 360
atatgcaaga ttgccagca gagatcaggt ccgagagttc ccttttttaa aaaaggagac 420
ttgcttaata aaagaagtct agccacgttt gtgtagagcg gctgtgctgt gctgggggtt 480

```


cacttttgag agagttctcc tctgagacct gatctctgga ggctgggcaa tcttgcaactt 540
 gagatggggc tgggtctgac tcagcactcc ttagtctgct cgcctctccc atggccccag 600
 cctggccaca cctgcttacg gggcactcct agatgcccac accataactt ccatgctagt 660
 ggactgtacc atatcagtg gaggctgcag caaggtggcc cctagagcca cgcaccagcc 720
 tgcacattgc ctctccatac ggcagccctt tatttgga aa cttcctaaat cactttgctg 780
 tgtgtgttta cacgggtgtg ttttgcttta cttgcccctga gagcacacgg gagtgcagca 840
 cacaccccaa cccacatcaa ctgccattaa agaaaagaaa tttcagccca gaatttcag 900
 tccagcaaaa ttaagcatca taagtgaagg agaaataaga tctttttcag acaagcaagt 960
 gctgagggaa tttggtatca ccagatctac cttacgagag ctctgaagg aagcactaaa 1020
 tatggaaaaga aaagatcatc acctgctact acaaaaacac actgaagtac acagtccaat 1080
 gatgctaaaa agcaagcaca tatgtaagtc tgcaaaaataa ccagctgaca gcatgacgac 1140
 aggataaaat ccacacatac cattactaac cttaaatgta aatgggctaa atgctcccat 1200
 tgaaagacac ggggcaagct gggtaaaagaa ccaagaccca ctggagtatg ccgtcttcaa 1260
 gcaaccatc tcacgtgcag tgccatacat aggtcaaaa taaaggaatg gagaaaaata 1320
 tttcaagcaa atggaaaaaca gaaaaaagg gttgcaactc cagtttctga caaacagac 1380
 tctaccaata aagataaaaa aagagaagga cattacaaag gtggtcctga cctttgataa 1440
 atctcattat tgcttgatac caacctgggc tatttgattt gcccaaacga ataggataat 1500
 ttgctgaggt tgtggagctt ctccccttca cagagtccct gatctccgaa aatttggtt 1560
 agatgtaagg ttgattttgc tgtacaactc cttttttgaa gttttactca tttccaacaa 1620
 ggaaggcaag ttttctgct tccattgaca aaggagagca ggcacctcct ttcctgagtt 1680
 tcagcttgct tctgacaggg aaggagcttt gagatttgaa tactggcctg ctgggttttg 1740
 gacgtgcatt gggcctgtgg tcccatttgt gttatttttc tgggaaattt cttccctttg 1800
 gagtgagaaa gcttacccaa tgctgtacc atcatcgta cttaaaagaa ctccatttta 1860
 agttcagga ctccctggca gaagagaccg tagccttgta tcagatcata aaggagaaga 1920
 gcaagaggtc cccggcaaac atccacagat ggccttgga ataagtcacc ttgctcacc 1980
 tgcaggaatg ccagtgaact tattgtgac atcttgagc tcagtacct catagtgtaa 2040
 cggcgctcag agatctgect gtgctgggac ttctgtact accattcct gagggcgat 2100
 gcttctgcag ggcctgtgac ttggtgcaca acttcagaca ccatcatctt gcagcagac 2160
 cgcacctca ctagccaggg tgtgatgac ttctcaagg ccaaggccac attcaaggct 2220
 tccgacttca ttgatgcgt tgtgctgac aagggtggct ctccgggatc ttaattcagg 2280
 aggtagaatg gagcttgaga tcaagtgtct gatcaagcct cagtgtatgg gcgctgttca 2340
 tctctgggtg ctgaagcagc caagagaccc aagtctgct ggctgcntct taggatatga 2400
 cagcagagcc agtggcctct actagatcct gtacaacctc acaaacacc cagacatcgg 2460
 gagtgtgccc agcctgtgat gcaagagtcc taatcctgaa gacattgaat gacctgtcat 2520
 tctgctgttt ttacaaaaa ggatcatgag gatcagagag gaaaagtcac ttgcccagg 2580
 tcacacagct gaacagtggg ggagttcaac tttgaccgtg ggctgtctga cccaagggtg 2640
 tatgcttgct tctctcccaa gagacaact tcttatcagg ctcaaatgaa tgaaaggagg 2700
 atgttaaagg taggatctct gaagcctgtg ccagtggaa cgcagctcat ggctggcacc 2760
 tgtgtttctca ttcttacctc attagagta aagtttattg agtttattga atttaagtat 2820
 ctttagtgag atcatatatt attagtaaga actgggacca aacagatttt ctgactctaa 2880
 aagagagatt ttcacagaaa cagatatata cctgtaagta tacagacacg catacacaca 2940
 tttctttact gctcataaaa attagtcctt attagaatgt gggatgtata aatgtaagag 3000
 aattttcatg ttaaaattga cagatacatt tttaaattgt cctaaaataa atttaattat 3060
 tttntttta gaattttcca ttattaatgt tttttttatg agaaactata taactttatt 3120
 gataatacat acaataacc tttgtttttc aaattgaaaa tacagtgtat tttgcaaata 3180
 actaagtcct aattttgtat taaaatttta aattttcaaa aaaaaaaa 3228

<210> 702

<211> 4894

<212> DNA

<213> Homo sapiens

<400> 702

tgtgatgtct ccgtacgtgt ggtggtgggt gageccaccg aggccagggt ggttccgggc 3300
 cggggcatct gcctggacct cgccatcctg gatagtgcct tcctgctgtc ccagggtggcc 3360
 ccatccctgt ttatgggctc cattgtccag ctcagccagt ctgtcactgc ctatatgggtg 3420
 tctgccgcag gcctgggtct ggtgcgccatt tactttgcta cacaggtagt atttgacaag 3480
 agcgacttgg ccaaatactc agcgtagaaa acttccagca cattgggggtg gagggcctgc 3540
 ctactggtt cccagctccc tgcctcctgtt agcccatgg ggctgccggg ctggccgccca 3600
 gtttctgttg ctgccaaagt aatgtggctc tctgctgccca cctgtgtctg ctgagggtgcg 3660
 tagctgcaca gctgggggct ggggcgtccc tctcctctct cccagctctc tagggctgcc 3720
 tgactggagg ccttccaagg ggggttcagt ctggacttat acagggaggc cagaagggt 3780
 ccatgcactg gaatgcgggg actctgcagg tggattacc aggctcaggg ttaacagcta 3840
 gcctcctagt tgagacacac cttagagaagg gttttggga gctgaataaa ctcagtcacc 3900
 tggtttccca tctctaagcc ccttaacctg cagcttcgtt taatgtagct cttgcatggg 3960
 agtttctagg atgaaacact ccaccatggg atttgaacat atgaaagtta tttgtagggg 4020
 aagagtcctg aggggcaaca cacaagaacc aggtcccctc agccacagc actgtctttt 4080
 tgctgatcca cccctctctt accttttatc aggatgtggc ctggttggtc tctgtttgcc 4140
 atcacagaga cacaggcatt taaatattta acttatttat ttaacaaagt agaagggaat 4200
 ccattgctag cttttctgtg ttggtgtcta atatttgggt aggggtgggg atccccaaca 4260
 atcaggtccc ctgagatagc tggtcattgg gctgatcatt gccagaatct tcttctcctg 4320
 gggctctggcc ccccaaatg cctaaccag gaccttggaa attctactca tcccaaatga 4380
 taattccaaa tgctgttacc caaggttagg gtgttggaagg aaggtagagg gtggggcttc 4440
 aggtctcaac ggcttcccta accaccctc ttctcttggc ccagcctggg tccccccact 4500
 tccactcccc tctactctct ctaggactgg gctgatgaag gcaactgccc aaatttcccc 4560
 tacccecaac tttccctac ccccaacttt cccaccagc tccacaacc tgtttgagc 4620
 tactgcagga ccagaagcac aaagtgcggg ttcccaagcc tttgtccatc tcagccccc 4680
 gagtatact gtgcttgggg aatctcacac agaaactcag gagcaccccc tgctgagct 4740
 aaggagggtc ttatctctca gggggggtt aagtgcctg tgcaataatg tctcttatt 4800
 tatttagcgg ggtgaatatt ttatactgta agtgagcaat cagagtataa tgtttatggt 4860
 gacaaaatta aaggctttct tatatgttta aaaa 4894

<210> 703

<211> 2904

<212> DNA

<213> Homo sapiens

<400> 703

gtctatgcct tcatgatcag tcttgggggc tgctgggct acctcctgcc tgccattgac 60
 tgggacacca gtgccctggc cccctacctg ggcaccagc aggagtgcct ctttggcctg 120
 ctcaccctca tcttctcac ctgcgtagca gccacactgc tgggtggctga ggaggcagcg 180
 ctgggccccca ccgagccagc agaagggtg tcggccccct ccttgtcgcc ccaactgctgt 240
 ccatgccggg cccgcttggc ttccgggaac ctgggcgccc tgcttccccg gctgcaccag 300
 ctgtgctgcc gcatgccccg caccctgcgc cggctcttcg tggctgagct gtgcagctgg 360
 atggcactca tgaccttcac gctgttttac acggatttcg tgggcgaggg gctgtaccag 420
 ggctgcccca gagctgagcc gggcaccgag gcccgagac actatgatga aggaaggcct 480
 ctggctgtct taggagtctg atcagagtcg ttgccccagt ttgacagaag gaaaggcgga 540
 gcttattcaa agtctagagg gagtgaggga gttaaggctg gatttcagat ctgcctgggt 600
 ccagcgcgag tgtgccctct gctcccccaa cgactttcca aataatctca ccagcgcctt 660
 ccagctcagg cgtcctagaa gcttcttgaa gcctatggcc agctgtcttt gtgttccctc 720
 tcaccgcct gtcctcacag ctgagactcc caggaaacct tcagactacc ttctctgcc 780
 ttcagcaagg ggcgttgccc acattctctg agggcgcttc gatgggcagc ctggggctgt 840
 tctgcaagt cgccatctcc ctggctctct ctctggtcat ggaccggctg gtgcagcgat 900
 tcggcactcg agcagtctat ttggccagtg tggcagcttt ccctgtggct gccggtgcca 960
 catgcctgtc ccacagtgtg gccgtgggtga cagcttcagc cgccctcacc gggttcacct 1020
 tctcagccct gcagatcctg ccctacacac tggcctccct ctaccaccgg gagaagcagg 1080

<210> 705
 <211> 6976
 <212> DNA
 <213> Homo sapiens

<400> 705

```

gaagctggac  cggcaccaaa  gggctggcag  aaatgggcgc  ctggctgatt  cctaggcagt  60
tggcggcagc  aaggaggaga  ggccgcagct  tctggagcag  agccgagacg  aagcagttct  120
ggagtgcctg  aacggccccc  tgagccctac  ccgcctggcc  cactatggtc  cagaggctgt  180
gggtgagccg  cctgctgcgg  caccgaaaag  ccagctctt  gctgggtcaac  ctgctaacct  240
ttggcctgga  ggtgtgtttg  gccgcaggca  tcacctatgt  gccgcctctg  ctgctggaag  300
tgggggtaga  ggagaagttc  atgaccatgg  tgctgggtga  gtcactacat  cctccttctt  360
tcctgttcca  gatacatgcc  acctggcatg  tgggacagga  gtacctctgc  cctgggagct  420
gcttggaggg  agaggtggtc  tgctgggaag  gcattgctgg  gcaggagggt  gacctggggc  480
tgagggggca  caccaagaga  aagaagagaa  taccaaggac  ataccccagt  cacctctgga  540
tcctgtgtcc  tgcacagagc  ctggctcata  ggagacactg  gagaaatgct  cctaaccttt  600
ggctagccct  tttataattt  atagcgatta  tctcatTTaa  tgcttacaac  caccatttga  660
ggtgatccat  tttacagaga  aggaagcaga  ggcttttaag  aggttaggta  agtcttagcc  720
aaagccaaat  agcagctgaa  cagtagagct  gggactccat  caaggtctcc  cagccggagc  780
ttgctcctac  ccctaggaca  aggggtggac  tcctgactct  gcagataaat  tctacaaaag  840
ccacagaagg  caagtagtaa  ccattgtgtg  acaaccctc  acccccagga  agagggggcc  900
ctgtgaggat  tgcaggctct  ggagtcacac  tgcttgttga  aacgctgcct  cttaccctcc  960
ctaggtctgc  gcctttgaat  aagtatcact  tmttagttgc  tccatgcctc  agtttgtcca  1020
tctgaaaatg  ggggcatctg  taatgcctgt  gttatgagga  gtaaattaca  gcatccctgt  1080
gaagacgtag  cacagtgtcg  agtacggaat  gttatttcca  tccttctcac  ggagcttggt  1140
tccccctccc  cttgcccttt  acttgtccca  gccattgact  catactactt  cccttcttgc  1200
aggcattggt  ccagtgtcgg  gccgtgtctg  tgtcccgctc  ctaggctcag  ccagtgacca  1260
ctggcgtaga  cgctatggcc  gccgcggcc  ctctcatctg  gcaactgtct  tgggcacact  1320
gctgagcctc  tttctcatcc  caagggccgg  ctggctagca  gggctgctgt  gcccggtacc  1380
caggccctcg  gagctggcac  tgctcactct  gggcggtggg  ctgctggact  tctgtggcca  1440
ggtgtgcttc  actccactgg  aggcctgtct  ctctgacctc  ttccgggacc  cggaccactg  1500
tcgccaggcc  tactctgtct  atgccttcat  gatcagctct  gggggctgcc  tgggctacct  1560
cctgcctgcc  attgactggg  acaccagtgc  cctggccccc  tacctgggca  cccaggagga  1620
gtgcctcttt  ggctgtctca  cctcatctct  cctcacctgc  gtagcagcca  cactgctggt  1680
ggctgaggag  gcagcgctgg  gcccaccga  gccagcagaa  gggctgtcgg  cccctcctt  1740
gtgcgccac  tgctgtccat  gccgggcccg  cttggctttc  cggaacctgg  gcgccctgct  1800
tccccggctg  caccagctgt  gctgccgcac  gccccgcacc  ctgcgccggc  tcttcgtggc  1860
tgagctgtgc  agctggatgg  cactcatgac  cttcacgctg  ttttacacgg  atttcgtggg  1920
cgaggggctg  taccaggcg  tgcccagagc  tgagccgggc  accgaggccc  ggagacacta  1980
tgatgaagg  aaggccttgg  cagccagcag  aggtctggtg  gggagccgcc  caccagagac  2040
gacactcggg  gctgtgtctg  ggctggtgcc  tctccatcct  ggccccgact  tctctgtcag  2100
gaaagtgggg  atggacccca  tctgcataca  cggttctca  tgggtgtgga  acatctctgc  2160
ttgcggtttc  aggaaggcct  ctggctgctc  taggagtctg  atcagagtct  ttgccccagt  2220
ttgacagaag  gaaaggcgga  gcttattcaa  agtctagagg  gagtggagga  gttaaaggctg  2280
gatttcagat  ctgcctgggt  ccagccgcag  tgtgccctct  gctcccccaa  cgactttcca  2340
aataatctca  ccagcgctct  ccagctcagg  cgtcctagaa  gcgtcttgaa  gcctatggcc  2400
agctgtcttt  gtgttccctc  tcaccgcct  gtcctcacag  ctgagactcc  caggaaaacct  2460
tcagactacc  ttctctgcc  ttcagcaagg  ggcgttgccc  acattctctg  agggtcagtg  2520
gaagaacct  gactccatt  gctagaggta  gaaaggggaa  ggggtgctgg  gagcagggct  2580
ggtccacagc  aggtctctg  cagcaggtac  ctgtggttcc  gccttctcat  ctccctgaga  2640
ctgctccgac  ccttccctcc  caggctctgt  ctgatggccc  ctctccctct  gcaggcgctt  2700
ggatgggcag  cctggggctg  ttctgcag  gcgccatctc  cctggctctt  tctctggtca  2760
tggaccggct  ggtgcagcga  ttccggcactc  gagcagtcta  tttggccagt  gtggcagctt  2820

```



```

cctccatggg atttgaacat atgaaagtta tttgtagggg aagagtcctg aggggcaaca 6120
cacaagaacc aggtcccctc agcccacagc actgtctttt tgctgatcca cccccctctt 6180
acctttttatc aggatgtggc ctgttggtcc ttctgttgcc atcacagaga cacaggcatt 6240
taaataattta acttatttat ttaacaaagt agaagggaat ccattgctag cttttctgtg 6300
ttgggtgtcta atatttgggt aggggtgggg atccccaaca atcagggtccc ctgagatagc 6360
tggtcattgg gctgatcatt gccagaatct tcttctcctg gggctctggc ccccaaaatg 6420
cctaaccag gaccttgga attctactca tcccaaatga taattccaaa tgctgttacc 6480
caaggttagg gtgttgaagg aaggtagagg gtggggcttc aggtctcaac ggcttcccta 6540
accacccctc ttctcttggc ccagcctggt tccccccact tccactcccc tctactctct 6600
ctaggactgg gctgatgaag gcaactgcca aaatttcccc taccaccaac tttcccctac 6660
cccaactttt cccaccagc tccacaacc tgtttgagc tactgcagga ccagaagcac 6720
aaagtgcggt ttccaagcc tttgtccatc tcagcccca gagtatact gtgcttgggg 6780
aatctcacac agaaactcag gagcaccccc tgctgagct aaggagggtc ttatctctca 6840
gggggggttt aagtgcggt tgcaataatg tegtcttatt tatttagcgg ggtgaatatt 6900
ttatactgta agtgagcaat cagagtataa tgtttatggt gacaaaatta aaggctttct 6960
tatatgttta aaaaaa
6976

```

<210> 706

<211> 123

<212> PRT

<213> Homo sapiens

<400> 706

```

Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu Val Phe
          5                      10                      15

```

```

Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val
          20                      25                      30

```

```

Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys
          35                      40                      45

```

```

Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly
          50                      55                      60

```

```

Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu
          65                      70                      75                      80

```

```

Tyr His Arg Glu Lys Gln Val Leu Ile Gly Gln Trp Val Glu Ser Gly
          85                      90                      95

```

```

Trp Glu Gly Trp Ser Gly Phe Leu Gly Gly Gln Leu Ala Gln Asn Leu
          100                      105                      110

```

```

Val Ser Gly Lys Gln Leu Trp Arg Met Leu Leu
          115                      120

```

<210> 707

<211> 150

<212> PRT

<213> Homo sapiens

009050"6226960

<400> 707

Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
 5 10 15

Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
 20 25 30

Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
 35 40 45

Glu Glu Lys Phe Met Thr Met Val Leu Gly Glu Ser Leu His Pro Pro
 50 55 60

Ser Phe Leu Phe Gln Ile His Ala Thr Trp His Val Gly Gln Glu Tyr
 65 70 75 80

Leu Cys Pro Gly Ser Cys Leu Glu Gly Glu Val Val Cys Trp Glu Gly
 85 90 95

Ile Ala Gly Gln Glu Gly Asp Pro Gly Leu Arg Gly His Thr Lys Arg
 100 105 110

Lys Lys Arg Ile Pro Arg Thr Tyr Pro Ser His Leu Trp Ile Pro Gly
 115 120 125

Pro Ala Gln Ser Leu Ala His Arg Arg His Trp Arg Asn Ala Pro Asn
 130 135 140

Leu Trp Leu Ala Leu Leu
 145 150

<210> 708

<211> 371

<212> PRT

<213> Homo sapiens

<400> 708

Met Leu Phe Pro Ser Phe Ser Arg Ser Leu Val Pro Leu Pro Leu Ala
 5 10 15

Leu Tyr Leu Ser Gln Pro Leu Thr His Thr Thr Ser Leu Leu Ala Gly
 20 25 30

Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala
 35 40 45

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
 50 55 60

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala
 65 70 75 80

Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu

[illegible]

370

<210> 709
 <211> 141
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(141)
 <223> n=A,T,C or G

<400> 709
 tacggcgtgg tgcggagggc ggtacccac aaataaacac nacaccccat cctatctgtg 60
 tccacanata aantgactca ttctctctct cgcatanccc actntcccct ngcgataccg 120
 taacnaancc cttccccctt t 141

<210> 710
 <211> 196
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(196)
 <223> n=A,T,C or G

<400> 710
 cnatccttcn cntacaccca tgangtccat gtcgcacgtc cacctcccct caaaacttgg 60
 gtcncatcc acccgctact ctcccentaa ncnataaccc cttttngcga atagacccca 120
 cettancaat nggttttttcn ttttttgtec ctnggnccgn gcgattcaan aaattgaagg 180
 cccanaaaaa ccccct 196

<210> 711
 <211> 177
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(177)
 <223> n=A,T,C or G

<400> 711
 ntacntcnct ccnaatgaaa ttcgaaanctc gggtacccgg gggnattccg attaggngcg 60
 tantctcgga tgtgcagtca caagtctttt gctaattctt ataattntcn ctaccctttc 120
 ttenacaata ctgctatctt anttnttctn tencctctct cccannttac taaccac 177

<210> 712
 <211> 185
 <212> DNA
 <213> Homo sapiens

003050"646550

<220>
 <221> misc_feature
 <222> (1)...(210)
 <223> n=A,T,C or G

<400> 719
 cancgctcgc ataacacgta ttttntgatn aagattctna ctgacccatn aantctacnt 60
 ctcaagctct tncanngtcc agtnaangga atgtgtatnn gtnggggatnc cacanaaaaa 120
 aganatntcg gncgcttcat tantcactct tcttaccan ntctctngat nncagnttg 180
 ancntgaacg cacactacng gatntctcca 210

<210> 720
 <211> 131
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(131)
 <223> n=A,T,C or G

<400> 720
 tccatcctaa tacgactcac tatagggctg ccaacctgcc atccactact gaggaagacc 60
 cgnanactta ggggtcact gcgagccacc ggccacaggt cgtatagggc aaagcacgng 120
 gaagcacccc t 131

<210> 721
 <211> 121
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(121)
 <223> n=A,T,C or G

<400> 721
 tccatcctaa tacgactcac tatagggccg ntgantnctg gcgaaaggct tacaattaag 60
 naggaaaaan ganccaacaa ctaaaaaaaaa nncggncgtg ncagcttnga tgactngtcc 120
 a 121

<210> 722
 <211> 246
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(246)
 <223> n=A,T,C or G

<400> 722
 anctggagtc gcgcgctgca gtcacattgt ggatccanaa aatcggcaca agctctcntg 60

gnttcntcga tatgaanaac actaatccca tgtngtntgn gtctccgtga ttcattccctc 120
gcacnggtcc ccntccnaac cnttgcatag gtgttatgtt gtantctccc cagtgcacaa 180
agattnacac tctctcantg tctganatat gcacgagttc attgtcctgt cnccgtnaac 240
atcaag 246

<210> 723

<211> 160

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(160)

<223> n=A,T,C or G

<400> 723

cctccggaaa atccaantag agtaantncn ctctaattccg gggnaattgg nggggttnnat 60
acgtctctct cccccagnt aggattnana aaaggntccc cagancaaaa nctccaaagt 120
gnatcnanta gccgtncctg ananccaacg cccctacgtc 160

<210> 724

<211> 156

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(156)

<223> n=A,T,C or G

<400> 724

tnanccnata tacaccaaatt tctgattcta aantcccacc caagggaaaa aagttgagaa 60
gagcctttcc acttttctac taataaaaaa atgcaccagc ccctaccann agtgnggaaa 120
acctccttag gcccttgntt ggaacaancg aaaatc 156

<210> 725

<211> 347

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(347)

<223> n=A,T,C or G

<400> 725

aganggttnt atncatgctg tactcgcgcg cctgcagtcg acactagtgg atccaaagaa 60
ttcggcacga gagacggtgc gcgatggacc gagggcccca gccgngagg cgccgccgcc 120
gagccgcggg ncagacgccc catcagtagc gtccgcaccg ggnagccgcg gntctcgccc 180
gagccgtggg cgcgcccagag gggcgggctc gcctcccgcc gtccctcgca gctctgccgg 240
gcccagagccc gcgcccgtcg cgccgccgnc ttgccgctcg gnccgcgcgg nccggnaaac 300
gcggctcgagg tctggatgng gcanngcccg cncctntcgc tgagcct 347

<210> 726

<211> 162
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(162)
 <223> n=A,T,C or G

<400> 726
 ttgggtgggt tgggtggggg naaatttncc catttggtg ggtttggggg ggnaaatact 60
 tccgccttt tngtnccca aaganacnaa gggggagtc cttnatagag gnagncgat 120
 ncntencaac nactngact ttgnccatgg ggagnaaggt gg 162

<210> 727
 <211> 120
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(120)
 <223> n=A,T,C or G

<400> 727
 gtgtgggtgg ggaattccat tgtggttggg ggnaaatctc cgcttgcca aagnacaggg 60
 ggggtcnctt anagnnagg gggttcctcc ccaccacttg ncttgnccat tngagnaag 120

<210> 728
 <211> 130
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(130)
 <223> n=A,T,C or G

<400> 728
 gacccactgc agcgtnaac ttagcttggg ccgagctcgg atccctagtc cgtgtggtgg 60
 aattccatgt gtcgagagag gggcaaatac nctccaanac ancncctca tgctcnacac 120
 atattcgcat 130

<210> 729
 <211> 182
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(182)
 <223> n=A,T,C or G

<400> 729

<400> 732

```

gcttgggtacc gagctnggat ccctagtaac ggccgccagt gtgctggaat tgggctttct 60
tcaatcagnt nacgagctgc atggtctgct aacattgtca taattgctgg catagattac 120
tgaaaaataaa gaaaaaaaaat tgaagctgcc tatcaagttt tgggtattatc aaaaacttcc 180
tacaagttat tttacttcaa ccatgttatt acaaataatt taatgaatac tttagagact 240
ttaattacaa aaaactgaga tagtaaaagc aagtaataaa agctgaaatt acttagctat 300
ttgataatta cataaattat tatggtccat tcaacttttc tagtgtttag tttatacacc 360
aggaagactt tcctattcta ctaacattta taaagtatgc taacctatta tttaaacgca 420
tccactatta ggattttatg gcctaaaacg tgatacagtt cagtatcttg atgtcaaaac 480
ttttaagca agtagggatt aagttcaagt gaatgtgatt ttctttcttc ccagtagggg 540
cttctgaata actcagnaaa gctcacttcc attatcttac tttataaaaa aatgctataa 600
gacagaatgg gccgacgtgg nggctccacc tgtatccacc tttggaggcg agnggcgaat 660

```

<210> 733

<211> 836

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(836)

<223> n=A,T,C or G

<400> 733

```

aattaatgac tttttttccg ccctgccaag ctagtgtgtc taaatataat gtaaagaaat 60
tagctactca ttttctgggc cacgaagggt cctaaaaatgg gaagaagtgg agatctgacc 120
ttgttagttc taaatacact aaactgggag tgccatggat ggctttcagg atgtcctgaa 180
tcctctataa ttgtatacaa aatcgtgagt ttttaaaaaac tgggttagag ctattgggtc 240
ctcagagtct caggcatctt agaccccaa aaagggttaag gactactgac ttaaccaatt 300
aggtttgagt ggcattggct ttgaagaaaa gcagaggaaa gatataatctt ataattctgg 360
gcaacaaaaa agtggatgtg tgccagcatc ttagagtaga atcctcttaa aaggatagca 420
ctgcatatga actagtaggt ttaaccagt gcatatttag gcgaagtagc tcatttttct 480
gttagaattc ttttttatct gggaatgggc aagcttttac agcttttacc ttgccaatga 540
atacctggaa ttttaaaaaat cttgttaggc atattgcca taaagttttt tttcctagat 600
catatattca gtaaataatg ttgtagcttt atttcaatcc cccaattcat tgaggggtga 660
aacaatttga atgggttgag tgtagaagct aagttatttc tgtagaggct aagggcattt 720
ataccaanat atggttagact tgnngntcct gttaaccatg ctgtanacaa taggaattac 780
tgtatatcca cattttaatt ttaacatctt ctgctttgnt gntgggttga gangga 836

```

<210> 734

<211> 694

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(694)

<223> n=A,T,C or G

<400> 734

```

nagtnctatt tncactaaac tngngagtgcc ttggatggct ttcaggatgt cctgaatcct 60
ctataattgt atacaaaatc gtgagttttt aaaaactggg ttagagctat tggttcctca 120
gagtctcagg catcttagac ccccaaaaag gttaaggact actgacttaa ccaattagggt 180
ttgagtggca ttggctttga agaaaagcag aggaaagata tattttataa ttctggggcaa 240

```

```

caaaaaagtg gatgtgtgcc agcatottag agtagaatcc tcttaaaagg atagcactgc 300
atatgaacta gtaggtttta accagtgcac atttaggcga agtagctcat ttttctgtta 360
gaattctttt ttatttgagg atgggcaagc ttttacagct ttaccttgc caatgaatac 420
ctggaattta aaaaatcttg ttaggcatac tgcccataaa gttttttttc ctagatcata 480
tattcagtaa atatgtttgt agctttatct caatccccc attcattgag ggttgaaaca 540
atttgaatgg tttgagtgtg gaagctaagt tatttctgtg gaggctaagg gcatttatac 600
caagatatgt tagacttgtg gttcctgtta accattgctg tagacaatag gaattactgt 660
atatccacat ttttaatttt aacatcattc tgctc 694

```

<210> 735

<211> 126

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(126)

<223> n=A,T,C or G

<400> 735

```

ncnttgaaac nggttgacca gacttcaggc ctgtgcgctc aatcgtggag aatctcgtgc 60
cgaattcggc acgagtctct ctctctctct ctctctctct ctctctctct ntctctctct 120
ctctctc 126

```

<210> 736

<211> 165

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(165)

<223> n=A,T,C or G

<400> 736

```

cagaagcctt taaaccggtt ngaccagact tcaggcctgt gcgctcaatc gtggagaatc 60
tcgtgccgaa ttcggcacga gtctctctct ctctctctct ctctctctct ctctctctct 120
ctctctctct ctctctctct ctctctctct ctctctctct ctctc 165

```

<210> 737

<211> 125

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(125)

<223> n=A,T,C or G

<400> 737

```

ggnagccctt ttaaccgttt gtccagactt caggcctgtg cgctcaatcg tggagaatct 60
cgtgcogaat tcggcacgag tctctctctc tctctctctc tctctctctc tctctntctc 120
tctct 125

```


<220>
 <221> misc_feature
 <222> (1)...(739)
 <223> n=A,T,C or G

<400> 742
 gntgtcnaaa aagcaggctg gtaccgggtcc ggaattcgcg gccgcgtcga cggcccttgg 60
 tgccactagt tctttcattc ttccccncca tcaatcagtg aacttttttag cctactcaaa 120
 gctttgctcc aatgcatagg atttatgatt gtgggggattt ccagataata taaatattca 180
 acatgaatat tttaaattaa ggcatgagac atttttccta actgagcata gccatgaacc 240
 tctcacgtct gttcctctgt gncagtttgt agcactgaat acagcagccc tcctaaaagt 300
 ccaggcagtg cacaggctctt gacatgatga agtgacgtgt tgctatgggtg attttgcagc 360
 tggccaaata gtcactgggtt gatttttacc agcaggagat ttttgcaaaa atttcctggg 420
 tgagagtga atcaaactcc tattttggtt ctctctgca agctgnagtt aanatggatt 480
 aatgagtact tttagattaa ttaactctga agagaaaatg ggagaaaagn gaggaagggtt 540
 gttggcagaa gtcattgctg gaatccttct gaagggagta ctgacttcac ttgcaaagac 600
 aagagactan aagacaatga agttaaactt ggctgtctn tcatatgata gatgcttgag 660
 agtacaggnt cagggaaatt ttaattctgn catacgcata ttggattatg tgggtcatgg 720
 ctttgtttgg cncctaacc 739

<210> 743
 <211> 610
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(610)
 <223> n=A,T,C or G

<400> 743
 ctgtccttat ttcttttagca aaaatttccc aagagaagaa ttgctgggat aatgcacatt 60
 taaatttttg atagacattc ccaaataatta tacctgtttt tgagaccttt aattcctgtt 120
 gtcaaattgc cctatatatg gagtaataaa cacgatttaa agaaatgagg actaaaaaaaa 180
 gattatatat aacccaacat aaaggcaacc tcttaggcgt tgacagaaac tgacaacttt 240
 ttatctgtgg gtgcgatcca ttataagtaa cctgagcacc ttattttttt tttttaaact 300
 ctaggttagga taccgaggt ccacaaattt ttcataagaa atattttttt tctgccctat 360
 gagattttta aaaatattat actgcttcaa ttgcatcaaa agaaatggac cctaatatct 420
 atgatgaagg atttggagtt agaagacctg agtttcaatt ttggcatggc tgtttgtcta 480
 gctctngat cttggacagg tcaattgact tggcttaatc ttctcatcca tttagnnggag 540
 acagcaccac tattcacagg actattgncn gaattaccag acaatagcat agngngaaaat 600
 ataangcett 610

<210> 744
 <211> 127
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(127)
 <223> n=A,T,C or G

<400> 744

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(738)

<223> n=A,T,C or G

<400> 747

```

gatatccccg gaattcgcgg ccgcgtcnac gaagcacaga cctgngccct gctctcatgg 60
ggcagactgc catttgtcat tnattactga aggaaagga tcctcagttt gcttgtggac 120
atttcaaatt tgaggtgaga gttggataag taagaataaa gctgctcttc aaagagatga 180
atatagaaaa agaaacaaga tacagncttg gcagtaaggc tgggaggaag gggaaaagg 240
aataaagaat gaaagagtga gaaatgtgag caggagctga acacagaaaa gttcagngac 300
agaagcanaa ggagggaaga agggaggagg gtccctttca cagaggctca cgaggatgct 360
ttatgngtgc catgcagtc atgttcagga tgtctgcttc ttanctctct acttttctaa 420
tanaaatttg gatacttact gatcctacat atgtaacagg gagagaagg gaatttcaaa 480
gcantaaatt gaaaaattgt tcacaatttc atttttttaa aaaagggagc taacagaaga 540
agagggttaat gtggttaatta taggatgnct cttgcgacac atgaatgnat ctggtatcat 600
ctgagtggga ggggagctgt cttcctgacc caaaaggatc ctttcgttan ccngnactta 660
ngtcccaaaa cctcaccacc ttggagaaat natttccttt tgggggtntc attaaancct 720
tttgncccc gcaaaagc                                     738

```

<210> 748

<211> 647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(647)

<223> n=A,T,C or G

<400> 748

```

ctntgtggcg gtggctgtct catttgggtg gacttttttg gtcgtaggaa cctggatatng 60
aggctcgagag taagacgggc tattagtagt cgcacgagag ttatttgtga aaacctggtt 120
agggcctctg tctccgctgc gctcgcctaa attggtatgg ctgcacttgg aaacacgggt 180
ctaacacgcg ttgttagcgc ccttgctagc atgtgaagga cactggccct accaagaaag 240
attcgagtcg ctccttcgcg tategttcac ggaggcgata tttactcttc ttactacggt 300
tacttcgaga ttgtctgtga agtttaagac tactaaaaag agtattaagc ctatcgggaa 360
ttagctagat cgacacgcta aaaccaaggg caatcggcgg aaatatagag gcaccaataa 420
tagggcctac agaaggcccg agggtttagac tcacgtttaa taccggccac gggagaaata 480
aaaagataaa gtatacatcg tttagcggtc ctcggaagcc ttcggcttta atgccaagga 540
gtcgaagca tcgtcggcga gtaataaact ccacgcgcc gagactatct acgacgccct 600
ccttaanatc cgtaaattac tcccggaaag agtatttagg cggctct                                     647

```

<210> 749

<211> 642

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(642)

<223> n=A,T,C or G

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(721)
<223> n=A,T,C or G

<400> 754
accggattng ttntctgagcg cgtgactgct aataaaaaag atggantgcc atctttttttt 60
ttnccttgct ttatatatcc agcagcaaaa caaaattggt ctgcngggct ataaaatttg 120
gcttgtagt cntgtacaca actcaggagt gtgacacagc taccagcttt cctcctaact 180
ctcaagggaa gaaaattcaa gttctgtcta ggctcactct gtaaagtggg aaacttgctg 240
gtttttagg ctttttttcc ctttctttcc ctctctcagc ttctccctgc ttctcagaan 300
atggagttgt gatgcctgca acttaccaaa tttatctatg aatcagattc cagtgggaga 360
cccctaaagc agagggagaa taaggagttc tccccatgat ggaaaatatc caaagacaag 420
gtttcatgga gcaaagaatt ctggctagat ttggtttgta agtggatccc tccccactgc 480
gtgtacactt tatctgtctc tttgcttctt cccaccctc tttcccagct ctctctctgt 540
ctctctcttg ntcccctgac ctttttttct tccantgca tacttttttn tttccctttt 600
ttaatcttct atantcttaa ncctaccaan gggccctcnt gannaatttn tcaccctga 660
ataggggatt cntangccc tgagaatttc nttatcanaa aaatattttt ttaaagcatt 720
a 721

<210> 755
<211> 721
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(721)
<223> n=A,T,C or G

<400> 755
accggattng ttntctgagcg cgtgactgct aataaaaaag atggantgcc atctttttttt 60
ttnccttgct ttatatatcc agcagcaaaa caaaattggt ctgcngggct ataaaatttg 120
gcttgtagt cntgtacaca actcaggagt gtgacacagc taccagcttt cctcctaact 180
ctcaagggaa gaaaattcaa gttctgtcta ggctcactct gtaaagtggg aaacttgctg 240
gtttttagg ctttttttcc ctttctttcc ctctctcagc ttctccctgc ttctcagaan 300
atggagttgt gatgcctgca acttaccaaa tttatctatg aatcagattc cagtgggaga 360
cccctaaagc agagggagaa taaggagttc tccccatgat ggaaaatatc caaagacaag 420
gtttcatgga gcaaagaatt ctggctagat ttggtttgta agtggatccc tccccactgc 480
gtgtacactt tatctgtctc tttgcttctt cccaccctc tttcccagct ctctctctgt 540
ctctctcttg ntcccctgac ctttttttct tccantgca tacttttttn tttccctttt 600
ttaatcttct atantcttaa ncctaccaan gggccctcnt gannaatttn tcaccctga 660
ataggggatt cntangccc tgagaatttc nttatcanaa aaatattttt ttaaagcatt 720
a 721

<210> 756
<211> 873
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(873)

<223> n=A,T,C or G

<400> 756

```

ggaagaatac agtaagtttg caaattaaaa tttctctatt tttctgttat ttattcattt 60
ggaaactgtc agcctgtctc tttcactttg ggcaagtgaag agcaaagacg tccagtccta 120
tcagcaatta ggctgaaagt caacgccaaag ctggcgggca agggctgggc tgagtagagg 180
ttccctaggc aggcaagaga gagactccca ctcgatactc ccagctcggc aactgcctga 240
atgccaatga gcaactcatta taaccgcgcc tattttatag gatttaattt tacacttcag 300
gcttaatcag tctgaaagtt aaactgacag tgttaagtta cggaatcaat gacatttagg 360
ctttatgact ttgtagctga atatctatgg gctatatctc cattctaaca gtgatatcct 420
gttcagaat ctcaattcttt ggtgatggca ctttctagtg gagcagtcac ggtaacagtc 480
cacaccatt accatgtggg tgcttttacag catactgacg gaaggactga ggagccaccg 540
gagcaggagt tctctcagg gaggaacgtg acaattccac agctgcctan gtatgggcac 600
ctgatgccaa cgaanaaccc aaagcgtctt ccttccaga tggaaagctgc cccacactgg 660
gctgacagca tctggagctg ctctggctca aatcccgaa tgcacacnt cctanccggg 720
gcgtttanag atcctcnggg ccagctaccg accacttttg acaaggggnt taggagcgat 780
aactagnctg gcgcgttaca cncggatgga acgtcttgga cttgagacct cttgggggan 840
atggcncccc caaataantt gggaaaantn ggg 873

```

<210> 757

<211> 782

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(782)

<223> n=A,T,C or G

<400> 757

```

ggccccctga gggatactct agagcggccg ccgactagtg agctcgtcga cgatatcccc 60
ggatttgaga ccaggagaca gctccagatg ctgtcagccc agtgctgggg gcaggcttcc 120
atctgtgaag tggagaggcg ctttgggctt cttcgttggc atcaggtgcc catacctagg 180
gcagctgtgg aagtgtcagc gtccctccctg agaggaaactc ctgctccggg ggctcctcag 240
tccctccgctc agtatgtctg aaagcaccca catggtaatg ggtgnggact ggtaccatga 300
ctgntccctt aaaaggtggc cttcccnaag aaaggagaat tcttggacna gggatttcac 360
ttgnttagaa atgggaaaaa ttaccatta gaattttcgn ttccaaggcn tnaagnocct 420
aaaggccttt gattcccgaa ccttaaccct gggcagttaa cctttcaaac gggataaacc 480
ctgangggga aatnaaate ctttaaaaaa ggggggggtt naaggagggc tctttggctt 540
tcaggcantt gccaacctgg gaaattcana ggggaagtnt ttttttttgc ctgcctaggg 600
aacctttact taaacnaacc cttgnccccc catttggggg tgactttcan cctaattgct 660
gaaaggaccg ggccgntttt gntttccttt gncccaaagg naaanaaacg ggtgccantt 720
cccangggat tanttcccga aaatttggnn aattttntt tгнаactttt tgggtttttt 780
cc 782

```

<210> 758

<211> 647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

005060" 005060" 005060"

<222> (1)...(647)

<223> n=A,T,C or G

<400> 758

```

ntttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatnga 60
gggaagagcg ccgtcgggtcc gagtacagta tggagtagta tagtcttcgc gccttctcgg 120
gcggcggggc tattctctcc aaaggcagag gtccctagtc gacctcgctc ccctaggtta 180
ggaacagccg tcgaatattt taggttcgtc gaggttttct tccgagctct acgcctaagt 240
agctccgcga gcaaagtatc ggtcattttc ccctatccat cactccccta agtacgcctc 300
attattccgg aaggcaagag gccagcattc ctccttagag tagagggtag gtacctccgt 360
cgcgtgccgc gaaagggcag agcttcgtgt cttccctccg cagcagctta acggtctacg 420
taggcgttct cgatcttttc acgggaatcg gggtcgggga gggcggcgga aaacgtcgac 480
gtctcgggtca ccgtcaccgc ccogaacaac tagcggcttt ccgctttcaa ctgaggaacc 540
ccgcacccct cattagcgct tacgaaatcg gggangtgat tgcgcccaatt cgtttagcctt 600
cgataattat tctctattag cggtcctatc tcgcgctttc gatttat 647

```

<210> 759

<211> 657

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(657)

<223> n=A,T,C or G

<400> 759

```

ctttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatnga 60
gggctctata gaaagcctct tgtcttttaga tacgggcttt ctggtccttc gttctggaag 120
tgtagtagta ggtactgcgg gaaggcgaag agtcctttca aggacgattt acttaagttg 180
gcttattcta tagttccttc gggacataag gtccgtacga tctatactgc gtgggaagct 240
gataggttgg gacttaaggc gaataagaag gaggcggcgg aggtcgcgat taccgcagag 300
atattattta cggcggccgc ggggtaccgc ggtcatgcgg aaattttctg aggttcttgg 360
attcctaaga tcgctcccggt cgagtatact agcgacgaac gtaagagtgc cctcacaaga 420
accggtacaa actcaagaag aagttcccat taagcatcgt aagaaacggg aggacgagga 480
cggtaagaag taatcggaga aaggatccta gtngttacga agaagcatcg ttnagctact 540
ttgcgctacc gtttatattt agacgtgttc cgtccttctc cgtgtttana aaaaagggtt 600
attccgacgg gagacttagg cgaatggagg gttccgcggg tganaatcgg ancgggg 657

```

<210> 760

<211> 644

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(644)

<223> n=A,T,C or G

<400> 760

```

ctttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatgna 60
ggaaaagaag taagcctcga agcctatctc cgaccgtatt tatttcgcag aagacggaac 120
tacggacgtc gttaaccccg agtagcccc gtaagaaagg actaaagcga atggaaaagt 180
cgggaattcc ggcggagggg cggcgattac tgaaaggagt aagagtaaga ctattgcgat 240

```

```

acttgaggcg ttccctctta aaaggcaccg gaaacactct attaaaaaac acccgaagaa 300
gaacaactca tgcgatcggc cgtgtgcagc cgtcaatagt aaagagagcc atgaaccatg 360
ccatccttag accaattagg atgaagaaga ggaggaagat gaggaccaa ccctaccac 420
tcggaaaacc ccgcacgagc ctccgaacaa aatccgggaa ttaaacggc ggcccacttc 480
cgcactctcg tagcgcggac cgaatagaaa accggaaact acagctaaag ggtcctttcc 540
ggcctgttat ctaccacccc gcaatccgat cctccccccc cctcgtccaa aaaccctaac 600
ctctgcggca acattagagc agaaggagag ggcgatccct tgan 644

```

```

<210> 761
<211> 647
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(647)
<223> n=A,T,C or G

```

```

<400> 761
ctttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatnga 60
ggcgggtact ctctgggata atcgggtataa gtgttgtaaa attgggggta agagaaagtt 120
tcattataag aagtggagc acgagccggg gtgttttagtc gttaatatta agaccggttt 180
ttgttgtagt tatatagctt gcgcgtgggg aggcaataag aaacattgcg tttcgaggcc 240
ggatgcgggg aacctctctc ggggtctaga gcgccgcac tgcaaaataa ggactactga 300
cgccgctcat aacgtactca acaatgagtc ggctgcatt aagatttcgg cgaagaaccg 360
tactgcgtct actgatagta tattgcattg atagcggcat gagctttatc acgtgtcgtt 420
ttcgggttgt aagaaggag ttaagtcgat cttcgaggaa gaagagacc caaataaaaa 480
atgactcaaa aaaacctaga agaaacacga cgaaaggaaa aagaacgtta aaactagtag 540
ctcttcggan gagtagcctt agtagggtta gtcctccgtg cgtactgtcc taagggttgg 600
atagcgcggg tgaatagacg gtcacgcgtc agaaggtaaa aanccgg 647

```

```

<210> 762
<211> 628
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(628)
<223> n=A,T,C or G

```

```

<400> 762
cattgtgttg gggctactga gccactttt ttccagattt tttgtaaaat tgtttcgcac 60
tgtgttccct ttattcgctt gtattaatat ttgcgtagtg gattaaacaa atacttggtg 120
ttgactgtca gtcttagagg actgactaga agtagttttc atttggggct caggaaatac 180
ctactttata tttctagcta attaggaag tcatttttca gttagggttg tgttttggtt 240
caggcactcg cttagctagat gacctaacat gctacttaat ttctgagtgt ttgtgtccat 300
ccctgtagga ttgttgccgg gttaaataaa attgtgtata tttgtaaagc atttacctca 360
gtgcccagac tgtgacagag tagattatta ggcttgctct tatttctgtg attaaattta 420
gtgtcagatt agcaacctat agctacttct aaagctgctg ctgctttctt tgtttagggt 480
taggaagaaa catgctggac agtttgccaa atgagagtta catgatgtgg cttgtgggaa 540
cattctaact tggaaacttg ccatttccag gactttgngg ttcanagatt tttggggata 600
gatgtaaggg ttaaaaaaaaa cngaaaac 628

```

<210> 763
 <211> 147
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(147)
 <223> n=A,T,C or G

<400> 763
 cattgtgttg gggcagagat aaataattcc tctgaaaagt gttttattgg aatttcaaatt 60
 gaaaagctaa ctggataact tacagcatgt ttctgccaat aatctcttan aacaggcctc 120
 ttttttttat gcacaccacc ttcnnggc 147

<210> 764
 <211> 146
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(146)
 <223> n=A,T,C or G

<400> 764
 cattgtgttg ggtatgtttt ttgaaggcag gtggacagga tttgctgatg ggtaaattggc 60
 agagtttaggg ggactgttag aacagagaaa ganatcatgg ggttgggttt gagtctgatg 120
 nnnaactggg gccgnntgct cagtat 146

<210> 765
 <211> 129
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(129)
 <223> n=A,T,C or G

<400> 765
 tncncgattc gntnctagcg tntacactna tgtcttggtta ccgagctcgg atccactagt 60
 ccagtgtggg nggaattcca ttgtgttggg gcaggaggng ctttgngtac ngtgcggtcg 120
 nagaggcgg 129

<210> 766
 <211> 175
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(175)
 <223> n=A,T,C or G

005060" 6225960

<211> 156
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(156)
 <223> n=A,T,C or G

<400> 771
 ttaaaaaanct ggnctccccg cggtggcggc cgctctagaa ctagtggatc cactagtcca 60
 gtgtggtgga attcgcggcc gcgtcgaccg cgagcggtcg cccttttttt tttttttttn 120
 ngtttttttg aanaattcat tgggtattta ttattc 156

<210> 772
 <211> 586
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(586)
 <223> n=A,T,C or G

<400> 772
 ncaanctggn ctccaccgcg gtggcgggcg ctctagacta gtggatccac tagtccagtg 60
 tgggtggaatt cgcgggccgcg tcgatcaca agtgctcaca agtcnngnat ttattttatc 120
 tccagatatg aaacttacc ccagctatgg tcttctattt gttatttaatt ttctaggcca 180
 attttttcca cttgaatgtc agtatttttaa ttcaaagtca ccttgtccaa ataccaagtc 240
 atcaacttac cctcaaatta tatcctcatt cagaaaatct acatctatta atggtagcta 300
 ttttatecct gccccctgct ttttcttttt atatttaatt aatttgntca tccagcaaat 360
 gcttattgag cagggtattgt aggctaaaca attctanact ttaaggggac acagnttgca 420
 aaacaaaatc ctgccttgna tggatactta tgnnatggng ggatacagac aatcaacata 480
 atgangngca tcatatataa tggttagnan aatgataagg gnttttggga aaaaaatgca 540
 cccanccaan anggattggg aagtggangg ganggtcang ggangg 586

<210> 773
 <211> 2983
 <212> DNA
 <213> Homo sapiens

<400> 773
 agagatagag tcttccctgg cattgcagga gagaatctga agggatgatg gatgcatcaa 60
 aagagctgca agttctccac attgacttct tgaatcagga caacgccgtt tctcaccaca 120
 catgggagtt ccaaacgagc agtcctgtgt tccggcgagg acaggtgttt cacctgcggc 180
 tgggtgctgaa ccagccccta caatcctacc accaactgaa actggaattc agcacagggc 240
 cgaatcctag catcgccaaa cacaccctgg tgggtgctga cccgaggacg ccctcagacc 300
 actacaactg gcaggcaacc cttcaaaatg agtctggcaa agaggtcaca gtggctgtca 360
 ccagttcccc caatgccatc ctgggcaagt accaactaaa cgtgaaaact ggaaaccaca 420
 tccttaagtc tgaagaaaac atcctatacc ttctcttcaa cccatggtgt aaagaggaca 480
 tggttttcat gcctgatgag gacgagcgca aagagtacat cctcaatgac acgggctgcc 540
 attacgtggg ggctgccaga agtatcaaat gcaaaccctg gaactttggt cagtttgaga 600
 aaaatgtcct ggactgctgc atttccctgc tgactgagag ctccctcaag cccacagata 660

ggaggggaccc cgtgctggtg tgcagggcca tgtgtgctat gatgagcttt gagaaaggcc 720
 agggcgctgct catttggaat tggactgggg actatgaagg tggcacagcc ccatacaagt 780
 ggacaggcag tgcgccgatc ctgcagcagt actacaacac gaagcaggct gtgtgctttg 840
 gccagtgtg ggtgtttgtc gggatcctga ctacagtgtc gagagcgttg ggcatcccag 900
 cacgcagtgt gacaggcttc gattcagctc acgacacaga aaggaacctc acggtggaca 960
 cctatgtgaa tgagaatggc aagaaaatca ccagtatgac ccacgactct gtctggaatt 1020
 tccatgtgtg gacggatgcc tggatgaagc gaccggatct gcccagggc tacgacggct 1080
 ggcaggctgt ggacgcaacg ccgcaggagc gaagccaggg tgtcttctgc tgtgggcat 1140
 caccactgac cgccatccgc aaaggtgaca tctttattgt ctatgacacc agattcgtct 1200
 tctcagaagt gaatggtgac aggtcactct ggttgggtgaa gatggtgaat gggcaggagg 1260
 agttacacgt aatttcaatg gagaccacaa gcatcgggaa aaacatcagc accaaggcag 1320
 tgggccaaga caggcggaga gatatacct atgagtacaa gtatccagaa ggctcctctg 1380
 aggagaggca ggtcatggat catgccttcc tcttctcag ttctgagagg gagcacagac 1440
 gacctgtaaa agagaacttt ctacacatgt cgggtacaatc agatgatgtg ctgctgggaa 1500
 actctgttaa tttcacctgt attcttaaaa ggaagaccgc tgccttacag aatgtcaaca 1560
 tcttgggctc ctttgaacta cagttgtaca ctggcaagaa gatggcaaaa ctgtgtgacc 1620
 tcaataagac ctgcgagatc caaggtcaag tatcagaagt gactctgacc ttggactcca 1680
 agacctacat caacagcctg gctatattag atgatgagcc agttatcaga ggtttcatca 1740
 ttgcggaaat tgtggagtct aaggaaatca tggcctctga agtattcacg tctttccagt 1800
 accctgagtt ctctatagag ttgcctaaca caggcagaat tggccagcta cttgtctgca 1860
 attgtatctt caagaatacc ctggccatcc ctttgactga cgtcaagttc tctttggaaa 1920
 gcctgggcat ctctcacta cagacctctg accatgggac ggtgcagcct ggtgagacca 1980
 tccaatccca aataaaatgc accccaataa aaactggacc caagaaattt atcgtcaagt 2040
 taagttccaa acaagtgaag gagattaatg ctacagaagt tgttctcacc accaagtagc 2100
 cttgtctgat gctgtggagc cttagttgag atttcagcat ttctacact gtgcttagct 2160
 ttacagattat ggatgattaa atttgatgac ttatatgagg gcagattcaa gagccagcag 2220
 gtcaaaaagg ccaacacaac cataagcagc cagaccaca aggccaggtc ctgtgctatc 2280
 acagggtcac ctcttttaca gttagaaaca ccagccgagg ccacagaatc ccatcccttt 2340
 cctgagtcac ggcctcaaaa atcagggccca ccattgtctc aattcaaatc catagatttc 2400
 gaagccacag agtctctccc tggagcagca gactatgggc agccagtgcc tgccacctgc 2460
 tgacgacctt tgagaagctg ccatactctc aggccatggg ttaccagcc ctgaaggcac 2520
 ctgtcaactg gagtgtctct tcagcactgg gatgggctg atagaagtgc attctcctcc 2580
 tattgcctcc attctcctct ctctatccct gaaatccagg aagtccctct cctggtgctc 2640
 caagcagttt gaagcccaat ctgcaaggac atttctcaag ggccatgtgg ttttgagac 2700
 aacctgtcc tcaggcctga actcaccata gagaccatg tcagcaaacg gtgaccagca 2760
 aatcctcttc ccttattcta aagctgcccc ttgggagact ccaggggagaa ggcattgctt 2820
 cctccctggt gtgaactctt tctttggtat tccatccact atcctggcaa ctcaaggctg 2880
 cttctgttaa ctgaagcctg ctcttcttg ttctgcctc cagagatttg ctcaaagtat 2940
 caataagctt taaattaaac tctacttcaa gaaaaaaaaa ccg 2983

<210> 774

<211> 3064

<212> DNA

<213> Homo sapiens

<400> 774

aattctaaaa atgcttttgc aagcttgcatt gcctgcaggc gcagcggccg ccagtgtgat 60
 ggatatctgc agaattcggc ttgcgctcag ctggaattcc gcagagatag agtcttccct 120
 ggcattgcag gagagaatct gaagggatga tggatgcac aaaagagctg caagttctcc 180
 acattgactt cttgaatcag gacaacgccg tttctcacca cacatgggag ttccaaacga 240
 gcagtccgtg gttccggcga ggacagggtg ttacactgcg gctggtgctg aaccagcccc 300
 tacaatccta ccaccaactg aaactggaat tcagcacagg gccgaatcct agcatcgcca 360
 aacacaccct ggtggtgctc gacccgagga cgcctcaga ccactacaac tggcaggcaa 420
 cccttcaaaa tgagtctggc aaagagggtc cagtggctgt caccagttcc ccaatgcca 480

Asp Thr Tyr Val Asn Glu Asn Gly Lys Lys Ile Thr Ser Met Thr His
 305 310 315 320
 Asp Ser Val Trp Asn Phe His Val Trp Thr Asp Ala Trp Met Lys Arg
 325 330 335
 Pro Asp Leu Pro Lys Gly Tyr Asp Gly Trp Gln Ala Val Asp Ala Thr
 340 345 350
 Pro Gln Glu Arg Ser Gln Gly Val Phe Cys Cys Gly Pro Ser Pro Leu
 355 360 365
 Thr Ala Ile Arg Lys Gly Asp Ile Phe Ile Val Tyr Asp Thr Arg Phe
 370 375 380
 Val Phe Ser Glu Val Asn Gly Asp Arg Leu Ile Trp Leu Val Lys Met
 385 390 395 400
 Val Asn Gly Gln Glu Glu Leu His Val Ile Ser Met Glu Thr Thr Ser
 405 410 415
 Ile Gly Lys Asn Ile Ser Thr Lys Ala Val Gly Gln Asp Arg Arg Arg
 420 425 430
 Asp Ile Thr Tyr Glu Tyr Lys Tyr Pro Glu Gly Ser Ser Glu Glu Arg
 435 440 445
 Gln Val Met Asp His Ala Phe Leu Leu Leu Ser Ser Glu Arg Glu His
 450 455 460
 Arg Arg Pro Val Lys Glu Asn Phe Leu His Met Ser Val Gln Ser Asp
 465 470 475 480
 Asp Val Leu Leu Gly Asn Ser Val Asn Phe Thr Val Ile Leu Lys Arg
 485 490 495
 Lys Thr Ala Ala Leu Gln Asn Val Asn Ile Leu Gly Ser Phe Glu Leu
 500 505 510
 Gln Leu Tyr Thr Gly Lys Lys Met Ala Lys Leu Cys Asp Leu Asn Lys
 515 520 525
 Thr Ser Gln Ile Gln Gly Gln Val Ser Glu Val Thr Leu Thr Leu Asp
 530 535 540
 Ser Lys Thr Tyr Ile Asn Ser Leu Ala Ile Leu Asp Asp Glu Pro Val
 545 550 555 560
 Ile Arg Gly Phe Ile Ile Ala Glu Ile Val Glu Ser Lys Glu Ile Met
 565 570 575
 Ala Ser Glu Val Phe Thr Ser Phe Gln Tyr Pro Glu Phe Ser Ile Glu
 580 585 590
 Leu Pro Asn Thr Gly Arg Ile Gly Gln Leu Leu Val Cys Asn Cys Ile

0057279-090600

Asn Asp Thr Gly Cys His Tyr Val Gly Ala Ala Arg Ser Ile Lys Cys
 165 170 175
 Lys Pro Trp Asn Phe Gly Gln Phe Glu Lys Asn Val Leu Asp Cys Cys
 180 185 190
 Ile Ser Leu Leu Thr Glu Ser Ser Leu Lys Pro Thr Asp Arg Arg Asp
 195 200 205
 Pro Val Leu Val Cys Arg Ala Met Cys Ala Met Met Ser Phe Glu Lys
 210 215 220
 Gly Gln Gly Val Leu Ile Gly Asn Trp Thr Gly Asp Tyr Glu Gly Gly
 225 230 235 240
 Thr Ala Pro Tyr Lys Trp Thr Gly Ser Ala Pro Ile Leu Gln Gln Tyr
 245 250 255
 Tyr Asn Thr Lys Gln Ala Val Cys Phe Gly Gln Cys Trp Val Phe Ala
 260 265 270
 Gly Ile Leu Thr Thr Val Leu Arg Ala Leu Gly Ile Pro Ala Arg Ser
 275 280 285
 Val Thr Gly Phe Asp Ser Ala His Asp Thr Glu Arg Asn Leu Thr Val
 290 295 300
 Asp Thr Tyr Val Asn Glu Asn Gly Glu Lys Ile Thr Ser Met Thr His
 305 310 315 320
 Asp Ser Val Trp Asn Phe His Val Trp Thr Asp Ala Trp Met Lys Arg
 325 330 335
 Pro Tyr Asp Gly Trp Gln Ala Val Asp Ala Thr Pro Gln Glu Arg Ser
 340 345 350
 Gln Gly Val Phe Cys Cys Gly Pro Ser Pro Leu Thr Ala Ile Arg Lys
 355 360 365
 Gly Asp Ile Phe Ile Val Tyr Asp Thr Arg Phe Val Phe Ser Glu Val
 370 375 380
 Asn Gly Asp Arg Leu Ile Trp Leu Val Lys Met Val Asn Gly Gln Glu
 385 390 395 400
 Glu Leu His Val Ile Ser Met Glu Thr Thr Ser Ile Gly Lys Asn Ile
 405 410 415
 Ser Thr Lys Ala Val Gly Gln Asp Arg Arg Arg Asp Ile Thr Tyr Glu
 420 425 430
 Tyr Lys Tyr Pro Glu Gly Ser Ser Glu Glu Arg Gln Val Met Asp His
 435 440 445

gtcacttagg aaaaggtgtc ctttcgggca gccgggtca gcatgaggaa cagaaggaat 60


```
<210> 778
<211> 1095
<212> PRT
<213> Homo sapiens
```

Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe

35 40 45
 Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala
 50 55 60
 Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp
 65 70 75 80
 Asn Tyr Lys Lys His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp
 85 90 95
 Ile Gln Phe Glu Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser
 100 105 110
 Cys Asp Thr Asp Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp
 115 120 125
 His Leu Lys Thr Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys
 130 135 140
 Asn Phe Ala Leu Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile
 145 150 155 160
 Tyr Ile Ala Gln Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His
 165 170 175
 Tyr Gly Leu Thr Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile
 180 185 190
 Ser Arg Ser Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp
 195 200 205
 Gly Met Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu
 210 215 220
 Gly Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
 225 230 235 240
 Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn
 245 250 255
 Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu
 260 265 270
 Glu Lys His Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly
 275 280 285
 Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly Lys Glu Thr Leu
 290 295 300
 Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val
 305 310 315 320
 Glu Gly Ser Gly Arg Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val

000000" 6225960

				325					330					335		
Glu	Asp	Ala	Pro	Thr	Ser	Ser	Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	
			340					345					350			
Leu	Pro	Arg	Thr	Val	Ser	Arg	Leu	Ser	Glu	Glu	Glu	Thr	Glu	Ser	Trp	
		355					360					365				
Ile	Lys	Trp	Leu	Lys	Glu	Ile	Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	
	370					375					380					
Ile	Lys	Met	Glu	Glu	Ala	Gly	Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	
385					390					395					400	
Tyr	Ala	Leu	Tyr	Lys	Ala	Phe	Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	
				405					410					415		
Trp	Asn	Gly	Gln	Leu	Lys	Leu	Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	
			420					425					430			
Ala	Asn	Asp	Glu	Ile	Phe	Thr	Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	
		435					440					445				
Leu	Gln	Glu	Val	Met	Phe	Thr	Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	
	450					455					460					
Val	Arg	Leu	Phe	Leu	Glu	Asn	Gly	Leu	Asn	Leu	Arg	Lys	Phe	Leu	Thr	
465					470					475					480	
His	Asp	Val	Leu	Thr	Glu	Leu	Phe	Ser	Asn	His	Phe	Ser	Thr	Leu	Val	
				485					490					495		
Tyr	Arg	Asn	Leu	Gln	Ile	Ala	Lys	Asn	Ser	Tyr	Asn	Asp	Ala	Leu	Leu	
			500					505					510			
Thr	Phe	Val	Trp	Lys	Leu	Val	Ala	Asn	Phe	Arg	Arg	Gly	Phe	Arg	Lys	
		515					520					525				
Glu	Asp	Arg	Asn	Gly	Arg	Asp	Glu	Met	Asp	Ile	Glu	Leu	His	Asp	Val	
	530					535					540					
Ser	Pro	Ile	Thr	Arg	His	Pro	Leu	Gln	Ala	Leu	Phe	Ile	Trp	Ala	Ile	
545					550					555					560	
Leu	Gln	Asn	Lys	Lys	Glu	Leu	Ser	Lys	Val	Ile	Trp	Glu	Gln	Thr	Arg	
				565					570					575		
Gly	Cys	Thr	Leu	Ala	Ala	Leu	Gly	Ala	Ser	Lys	Leu	Leu	Lys	Thr	Leu	
			580					585					590			
Ala	Lys	Val	Lys	Asn	Asp	Ile	Asn	Ala	Ala	Gly	Glu	Ser	Glu	Glu	Leu	
		595					600						605			
Ala	Asn	Glu	Tyr	Glu	Thr	Arg	Ala	Val	Glu	Leu	Phe	Thr	Glu	Cys	Tyr	

610 615 620
 Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu
 625 630 635 640
 Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp
 645 650 655
 Gln His Phe Thr Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln
 660 665 670
 Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu
 675 680 685
 Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg
 690 695 700
 Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
 705 710 715 720
 Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe Tyr
 725 730 735
 Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp Phe His
 740 745 750
 Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu Val Phe Val
 755 760 765
 Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn Gly Val Asn Tyr
 770 775 780
 Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu Gly Leu Phe Tyr Phe
 785 790 795 800
 Ile Ala Gly Ile Val Phe Arg Leu His Ser Ser Asn Lys Ser Ser Leu
 805 810 815
 Tyr Ser Gly Arg Val Ile Phe Cys Leu Asp Tyr Ile Ile Phe Thr Leu
 820 825 830
 Arg Leu Ile His Ile Phe Thr Val Ser Arg Asn Leu Gly Pro Lys Ile
 835 840 845
 Ile Met Leu Gln Arg Met Leu Ile Asp Val Phe Phe Phe Leu Phe Leu
 850 855 860
 Phe Ala Val Trp Met Val Ala Phe Gly Val Ala Arg Gln Gly Ile Leu
 865 870 875 880
 Arg Gln Asn Glu Gln Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr
 885 890 895
 Glu Pro Tyr Leu Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly

000000" 6225960

```
<210> 779
<211> 3639
<212> DNA
<213> Homo sapiens
```

<400> 779						
gattacgcaa	gctat ttagg	tgacactata	gaatwctcag	cttgcacaa	gcttggtacc	60
gagctcggat	ccctagtaac	ggccgccagt	gtgctggaat	tcgcccttgc	agccgggctc	120
agcatgagga	acagaaggaa	tgacactctg	gacagcacc	ggaccctgta	ctccagcgcg	180
tctcggagca	cagacttgct	ttacagtga	agcgacttgg	tgaattttat	tcaagcaa	240
tttaagaaac	gagaatgtgt	cttctttacc	aaagattcca	aggccacgga	gaatgtgtgc	300
aagtgtggct	atgccagag	ccagcacatg	gaaggcacc	agatcaacca	aagtgagaaa	360
tggaactaca	agaaacacac	caaggaattt	cctaccgacg	cctttgggga	tattcagttt	420
gagacactgg	ggaagaaagg	gaagtatata	cgtctgtcct	gcgacacgga	cgcggaatc	480


```
<220>  
<221> VARIANT  
<222> (1)...(1095)  
<223> Xaa = Any Amino Acid
```

<400>	780																		
Met	Arg	Asn	Arg	Arg	Asn	Asp	Thr	Leu	Asp	Ser	Thr	Arg	Thr	Leu	Tyr				
				5					10					15					
Ser	Ser	Ala	Ser	Arg	Ser	Thr	Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu				
			20					25					30						
Val	Asn	Phe	Ile	Gln	Ala	Asn	Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe				
		35					40					45							
Thr	Lys	Asp	Ser	Lys	Ala	Thr	Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala				
	50					55					60								
Gln	Ser	Gln	His	Met	Glu	Gly	Thr	Gln	Ile	Asn	Gln	Ser	Glu	Lys	Trp				
65					70					75					80				
Asn	Tyr	Lys	Lys	His	Thr	Lys	Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp				
				85					90					95					
Ile	Gln	Phe	Glu	Thr	Leu	Gly	Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser				
			100					105					110						
Cys	Asp	Thr	Asp	Ala	Glu	Ile	Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp				
		115					120					125							
His	Leu	Lys	Thr	Pro	Asn	Leu	Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys				
	130					135					140								
Asn	Phe	Ala	Leu	Lys	Pro	Arg	Met	Arg	Lys	Ile	Phe	Ser	Arg	Leu	Ile				
145					150					155					160				
Tyr	Ile	Ala	Gln	Ser	Lys	Gly	Ala	Trp	Ile	Leu	Thr	Gly	Gly	Thr	His				
			165						170					175					
Tyr	Gly	Leu	Met	Lys	Tyr	Ile	Gly	Glu	Val	Val	Arg	Asp	Asn	Thr	Ile				
			180					185					190						
Ser	Arg	Ser	Ser	Glu	Glu	Asn	Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp				
		195					200					205							
Gly	Met	Val	Ser	Asn	Arg	Asp	Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu				
	210					215					220								
Gly	Tyr	Phe	Leu	Ala	Gln	Tyr	Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro				

515 520 525
 Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu Leu His Asp Val
 530 535 540
 Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu Phe Ile Trp Ala Ile
 545 550 555 560
 Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg
 565 570 575
 Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu
 580 585 590
 Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu
 595 600 605
 Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr
 610 615 620
 Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu
 625 630 635 640
 Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp
 645 650 655
 Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln
 660 665 670
 Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu
 675 680 685
 Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg
 690 695 700
 Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
 705 710 715 720
 Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe Tyr
 725 730 735
 Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp Phe His
 740 745 750
 Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu Val Phe Val
 755 760 765
 Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn Gly Val Asn Tyr
 770 775 780
 Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu Gly Leu Phe Tyr Phe
 785 790 795 800
 Ile Ala Gly Ile Val Phe Arg Leu His Ser Ser Asn Lys Ser Ser Leu

805 810 815
 Tyr Ser Gly Arg Val Ile Phe Cys Leu Asp Tyr Ile Ile Phe Thr Leu
 820 825 830
 Arg Leu Ile His Ile Phe Thr Val Ser Arg Asn Leu Gly Pro Lys Ile
 835 840 845
 Ile Met Leu Gln Arg Met Leu Ile Asp Val Phe Phe Phe Leu Phe Leu
 850 855 860
 Phe Ala Xaa Trp Met Val Ala Phe Gly Val Ala Arg Gln Gly Ile Leu
 865 870 875 880
 Arg Gln Asn Glu Gln Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr
 885 890 895
 Glu Pro Tyr Leu Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly
 900 905 910
 Thr Thr Tyr Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys
 915 920 925
 Pro Leu Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu
 930 935 940
 Trp Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
 945 950 955 960
 Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly Thr
 965 970 975
 Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr Phe Leu
 980 985 990
 Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro Phe Ile Val
 995 1000 1005
 Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe Lys Cys Cys Cys
 1010 1015 1020
 Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys Phe Lys Asn Glu Asp
 1025 1030 1035 1040
 Asn Glu Thr Leu Ala Trp Glu Gly Val Met Lys Glu Asn Tyr Leu Val
 1045 1050 1055
 Lys Ile Asn Thr Lys Ala Asn Asp Thr Ser Glu Glu Met Arg His Arg
 1060 1065 1070
 Phe Arg Gln Leu Asp Thr Lys Leu Asn Asp Leu Lys Gly Leu Leu Lys
 1075 1080 1085
 Glu Ile Ala Asn Lys Ile Lys

tcgcagccct ggcaggcggc actgggtcatg gaaaacgaat tgttctgctc g 51

<210> 794
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 794
 atcagcattg cttcgcagtg ccctaccgcg gggaactctt gcctc 45

<210> 795
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 795
 tccgtgtccg agtctgacac catccggagc atcagcattg cttcg 45

<210> 796
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 796
 atcaagttgg acgaatccgt gtccgagtct gacaccatcc ggagc 45

<210> 797
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 797
 aacgacctca tgctcatcaa gttggacgaa tccgtgtccg agtct 45

<210> 798
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 798
 agacccttgc tcgctaacga cctcatgctc atcaagttgg acgaa 45

<210> 799
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 799
 Glu Pro Gly Ser Gln Met Val Glu Ala Ser Leu Ser Val Arg His
 5 10 15

<210> 800
 <211> 15

<400> 810
Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly Asn Ser Cys Leu

```
<210> 811
<211> 15
<212> PRT
<213> Homo sapiens
```

```
<210> 812
<211> 15
<212> PRT
<213> Homo sapiens
```

```
<210> 813
<211> 15
<212> PRT
<213> Homo sapiens
```

```
<210> 814
<211> 15
<212> PRT
<213> Homo sapiens
```

```
<210> 815
<211> 35
<212> DNA
<213> Artificial Sequence
```

```
<400> 815
ggaccagcat atgaggaaca gaaggaatga cactc
```

<210> 816
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 816
 ccgctcagagt ccaccccaag cttcacagg

29

<210> 817
 <211> 1959
 <212> DNA
 <213> Homo sapiens

<400> 817
 atgaggaaca gaaggaatga cactctggac agcacccgga ccctgtactc cagcgcgtct 60
 cggagcacag acttgtctta cagtgaagc gacttggtga attttattca agcaaatttt 120
 aagaaacgag aatgtgtctt ctttaccaaa gattccaagg ccacggagaa tgtgtgcaag 180
 tgtggctatg cccagagcca gcacatggaa ggcacccaga tcaaccaaag tgagaaatgg 240
 aactacaaga aacacaccaa ggaatttcct accgacgcct ttggggatat tcagtttgag 300
 acactgggga agaaagggaa gtatatacgt ctgtcctgcg acacggacgc ggaaatcctt 360
 tacgagctgc tgaccagca ctggcacctg aaaacaccca acctggtcat ttctgtgacc 420
 gggggcgcca agaacttcgc cctgaagcgc cgcctgcgca agatcttcag ccggctcatc 480
 tacatcgcg agtccaaagg tgcttggtt ctcacgggag gcacccatta tggcctgatg 540
 aagtacatcg gggaggtggt gagagataac accatcagca ggagttcaga ggagaatatt 600
 gtggccattg gcatagcagc ttggggcatg gtctccaacc gggacacct catcaggaat 660
 tgcgatctg agggctattt tttagcccag taccttatgg atgacttcac aagagatcca 720
 ctgtatatcc tggacaacaa ccacacacat ttgtctctcg tggacaatgg ctgtcatgga 780
 catccactg tcgaagcaaa gctccggaat cagctagaga agtatatctc tgagcgcact 840
 attcaagatt ccaactatgg tggcaagatc ccctatttgt gttttgcca aggaggtgga 900
 aaagagactt tgaaagccat caatacctcc atcaaaaata aaattccttg tgtggtggtg 960
 gaaggctcgg gccagatcgc tgatgtgatc gctagcctgg tggaggtgga ggtgcccctg 1020
 acatcttctg ccgtcaagga gaagctggtg cgctttttac cccgcacggt gtcccggctg 1080
 cctgaggagg agactgagag ttggatcaaa tggctcaaag aaattctcga atgttctcac 1140
 ctattaacag ttattaaaat ggaagaagct ggggatgaaa ttgtgagcaa tgccatctcc 1200
 tacgctctat acaaagcctt cagcaccagt gagcaagaca aggataactg gaatgggcag 1260
 ctgaagcttc tgctggagtg gaaccagctg gacttagcca atgatgagat tttaccaat 1320
 gaccgccgat gggagtctgc tgaccttcaa gaagtcatgt ttacggctct cataaaggac 1380
 agacccaagt ttgtccgcct ctttctggag aatggcttga acctacggaa gtttctcacc 1440
 catgatgtcc tcaactgaact cttctccaac cacttcagca cgcttggtga ccggaatctg 1500
 cagatcgcca agaattccta taatgatgcc ctctcactg ttgtctggaa actggttgcg 1560
 aacttccgaa gaggtctccg gaaggaagac agaaatggcc gggacgagat ggacatagaa 1620
 ctccacgacg tgtctcctat tactcggcac cccctgcaag ctctcttcat ctgggccatt 1680
 cttcagaata agaaggaact ctccaaagtc atttgggagc agaccagggg ctgcactctg 1740
 gcagccctgg gagccagcaa gcttctgaag actctggcca aagtgaagaa cgacatcaat 1800
 gctgctgggg agtccgagga gctggctaat gagtacgaga cccgggctgt tgagctgttc 1860
 actgagtgtt acagcagcga tgaagacttg gcagaacagc tgctggtcta ttctgtgaa 1920
 gcttgggggtg gactcgagca ccaccaccac caccactga 1959

<210> 818
 <211> 652

<212> PRT

<213> Homo sapiens

<400> 818

```

Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr
                    5                      10                      15

Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu
                20                      25                      30

Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe
                35                      40                      45

Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala
                50                      55                      60

Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp
                65                      70                      75                      80

Asn Tyr Lys Lys His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp
                85                      90                      95

Ile Gln Phe Glu Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser
                100                     105                     110

Cys Asp Thr Asp Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp
                115                     120                     125

His Leu Lys Thr Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys
                130                     135                     140

Asn Phe Ala Leu Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile
                145                     150                     155                     160

Tyr Ile Ala Gln Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His
                165                     170                     175

Tyr Gly Leu Met Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile
                180                     185                     190

Ser Arg Ser Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp
                195                     200                     205

Gly Met Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu
                210                     215                     220

Gly Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
                225                     230                     235                     240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn
                245                     250                     255

Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu
                260                     265                     270

```

003060"0223360

Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg
565 570 575

Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu
580 585 590

Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu
595 600 605

Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr
610 615 620

Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu
625 630 635 640

Ala Trp Gly Gly Leu Glu His His His His His His
645 650

<210> 819

<211> 132

<212> PRT

<213> Homo sapien

<400> 819

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1 5 10 15

Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20 25 30

Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35 40 45

Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50 55 60

Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65 70 75 80

Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85 90 95

Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
100 105 110

Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115 120 125

Gly Pro Pro Ala
130

<210> 820

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

005060"6225960

<223> PCR primer

<400> 820

ggggaattca tgatccggga gaaatttgcc cactgc

36

<210> 821

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 821

gggctcgagt caggagtttg agaccagcct ggc

33

<210> 822

<211> 675

<212> DNA

<213> Homo sapiens

<400> 822

atgcatcacc	atcaccatca	cacggccgcg	tccgataact	tccagctgtc	ccaggggtggg	60
cagggattcg	ccatttccgat	cgggcaggcg	atggcgatcg	cgggccagat	caagcttccc	120
accgttcata	tcgggcctac	cgccttcctc	ggcttggtg	ttgtcgacaa	caacggcaac	180
ggcgacgag	tccaacgcgt	ggtcgggagc	gctccggcgg	caagtctcgg	catctccacc	240
ggcgacgtga	tcaccgcggg	cgacggcgct	ccgatcaact	cggccaccgc	gatggcggac	300
gcgcttaacg	ggcatcatcc	cggtgacgtc	atctcgggtga	cctggcaaac	caagtcgggc	360
ggcacgcgta	cagggaaacgt	gacattggcc	gagggacccc	cggccgaatt	catgatccgg	420
gagaaatttg	cccactgcac	cgtgctaacc	attgcacaca	gattgaacac	cattattgac	480
agcgacaaga	taatggtttt	agattcagga	agactgaaag	aatatgatga	gccgtatggt	540
ttgctgcaaa	ataaagagag	cctattttac	aagatgggtgc	aacaactggg	caaggcagaa	600
gccgctgccc	tcactgaaac	agcaaaacag	agatgggggtt	tcaccatggt	ggccaggctg	660
gtctcaaact	cctga					675

<210> 823

<211> 291

<212> DNA

<213> Homo sapiens

<400> 823

atggggatcc	gggagaaatt	tgcccactgc	accgtgctaa	ccattgcaca	cagattgaac	60
accattattg	acagcgacaa	gataatgggt	ttagattcag	gaagactgaa	agaatatgat	120
gagccgtatg	ttttgctgca	aaataaagag	agcctatttt	acaagatggg	gcaacaactg	180
ggcaaggcag	aagccgctgc	cctcactgaa	acagcaaaac	agagatgggg	tttcaccatg	240
ttggccaggc	tgggtctcaa	ctccctcgag	caccaccacc	accaccactg	a	291

<210> 824

<211> 1074

<212> DNA

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Ile Arg Glu Lys Phe Ala
130 135 140

<210> 827

<211> 96

<212> PRT

<213> Homo sapiens

<400> 827

Met Gly Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr Ile Ala

His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met Val Leu Asp
20 25 30

Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val Leu Leu Gln Asn
 35 40 45

Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln Leu Gly Lys Ala Glu
 50 55 60

Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Arg Trp Gly Phe Thr Met
 65 70 75 80

Leu Ala Arg Leu Val Ser Asn Ser Leu Glu His His His His His His
 85 90 95

<210> 828

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 828

cgcccatggg gatccgggag aaatttgccc actgc

35

<210> 829

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 829

cgcctcgagg gagtttgaga ccagcctggc caaca

35

<210> 830

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 830

gcatggacca tatgtcagcc attgagaggg tgtcagag

38

<210> 831

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

009060" 6225960

<223> PCR primer

<400> 831

ccgctcgaga ataaggaaaa tgaagacaat ccag

34

<210> 832

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 832

gttgaattca tgcacggggcc ccagggtg

27

<210> 833

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 833

cccctcgagt cactatgggtc tgcctcttga

30

<210> 834

<211> 915

<212> DNA

<213> Homo sapiens

<400> 834

atgcacacacc	atcacatca	cacggccgcg	tccgataact	tccagctgtc	ccaggggtggg	60
cagggattcg	ccattccgat	cgggcaggcg	atggcgatcg	cgggccagat	caagcttccc	120
accgttcata	tcgggcctac	cgccttcctc	ggcttggtg	ttgtcgaaa	caacggcaac	180
ggcgacagag	tccaacgcgt	ggtcgggagc	gctccggcgg	caagtctcgg	catctccacc	240
ggcgacgtga	tcaccgcggt	cgacggcgct	ccgatcaact	cggccaccgc	gatggcggac	300
gcgcttaacg	ggcatcatcc	cgggtgacgtc	atctcggtga	cctggcaaac	caagtcgggc	360
ggcacgcgta	cagggaaacgt	gacattggcc	gagggacccc	cggccgaatt	catgcacggg	420
cccaggtgc	tggcacgctg	ctcogagtgt	gcttgctctg	ccttggtctc	cacctctgcg	480
gggtgctgc	tggagggggg	ggaccggcca	ccaaccttac	ccagtcaagg	aagtggatgg	540
ccatgttccc	acagcctgag	tggctgccac	ctgatggctg	atggagcaaa	ggccttagga	600
aaagcagatg	gcccttggtc	ctaccttttt	gttagaagaa	ctgatgttcc	atgtcctgca	660
gcgagtggag	ttggtggctg	tgcctccagc	tcctggcgcg	ccctcgcaga	ggtgactggg	720
tgctctttgg	gccctcttgg	ccttgcccag	catgcacaag	cctcagtgct	actactgtgc	780
tacaaatgga	gccatatagg	ggaaacgagc	agccatctca	ggagcaagg	gtatgctgcc	840
tttgggggct	ccagtccttg	cctcaagggt	cttatgtcac	tgtgggcttc	ttggttgtca	900
agaggcagac	catag					915

<210> 835

<211> 304
 <212> PRT
 <213> Homo sapiens

<400> 835

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
      5                      10                      15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
      20                      25                      30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
      35                      40                      45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
      50                      55                      60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
      65                      70                      75                      80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
      85                      90                      95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
      100                     105                     110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
      115                     120                     125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met His Gly Pro Gln Val Leu
      130                     135                     140

Ala Arg Cys Ser Glu Cys Ala Cys Pro Ala Leu Ala Ala Thr Ser Ala
      145                     150                     155                     160

Gly Val Arg Leu Glu Gly Val Asp Arg Pro Pro Thr Leu Pro Ser Gln
      165                     170                     175

Gly Ser Gly Trp Pro Cys Ser His Ser Leu Ser Gly Cys His Leu Met
      180                     185                     190

Ala Asp Gly Ala Lys Ala Leu Gly Lys Ala Asp Gly Pro Trp Pro Tyr
      195                     200                     205

Leu Phe Val Arg Arg Thr Asp Val Pro Cys Pro Ala Ala Ser Glu Val
      210                     215                     220

Gly Gly Cys Ala Pro Ser Ser Trp Arg Ala Leu Ala Glu Val Thr Gly
      225                     230                     235                     240

Cys Ser Leu Gly Pro Leu Gly Leu Ala Gln His Ala Gln Ala Ser Val
      245                     250                     255

Leu Leu Leu Cys Tyr Lys Trp Ser His Ile Gly Glu Thr Ser Ser His

```

009060"6225960

```
<210> 836
<211> 24
<212> DNA
<213> Artificial Sequence
```

<400> 836
cgaagtccacg tggaggccag cctc 24

<220>
<223> PCR primer

```
<210> 838
<211> 166
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> VARIANT  
<222> (1) ... (166)  
<223> Xaa = Any Amino Acid
```

```

<400> 838
Met Gly His His His His His His Val Glu Ala Ser Leu Ser Val Arg
 1             5             10             15
His Pro Glu Tyr Asn Arg Pro Leu Leu Ala Asn Asp Leu Met Leu Ile
          20             25             30
Lys Leu Asp Glu Ser Val Ser Glu Ser Asp Thr Ile Arg Ser Ile Ser
          35             40             45
Ile Ala Ser Gln Cys Pro Thr Ala Gly Asn Ser Cys Leu Val Ser Gly
      50             55             60
Trp Gly Leu Leu Ala Asn Gly Arg Met Pro Thr Val Leu Gln Cys Val

```

<220>
<223> PCR primer

<400> 841
ctatagaatt cattaccaa aagctgggct ccagc

35

<210> 842
<211> 241
<212> PRT
<213> Homo sapiens

<400> 842
Met Gln His His His His His His Leu Arg Val Pro Glu Pro Arg Pro
1 5 10 15
Gly Glu Ala Lys Ala Glu Gly Ala Ala Pro Pro Thr Pro Ser Lys Pro
20 25 30
Leu Thr Ser Phe Leu Ile Gln Asp Ile Leu Arg Asp Gly Ala Gln Arg
35 40 45
Gln Gly Gly Arg Thr Ser Ser Gln Arg Gln Arg Asp Pro Glu Pro Glu
50 55 60
Pro Glu Pro Glu Pro Glu Gly Gly Arg Ser Arg Ala Gly Ala Gln Asn
65 70 75 80
Asp Gln Leu Ser Thr Gly Pro Arg Ala Ala Pro Glu Glu Ala Glu Thr
85 90 95
Leu Ala Glu Thr Glu Pro Glu Arg His Leu Gly Ser Tyr Leu Leu Asp
100 105 110
Ser Glu Asn Thr Ser Gly Ala Leu Pro Arg Leu Pro Gln Thr Pro Lys
115 120 125
Gln Pro Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln Val Ile
130 135 140
Glu Leu Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala Pro Glu
145 150 155 160
Arg Ala His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln Val Lys
165 170 175
Ile Trp Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln Leu Ser
180 185 190
Ser Glu Leu Gly Asp Leu Glu Lys His Ser Ser Leu Pro Ala Leu Lys
195 200 205
Glu Glu Ala Phe Ser Arg Ala Ser Leu Val Ser Val Tyr Asn Ser Tyr
210 215 220
Pro Tyr Tyr Pro Tyr Leu Tyr Cys Val Gly Ser Trp Ser Pro Ala Phe
225 230 235 240
Trp

<210> 843
<211> 729
<212> DNA
<213> Homo sapiens

<400> 843
atgcagcatc accaccatca ccacctcagg gttccggagc cgcgggcccg ggaggcgaaa 60
gcggaggggg ccgcgccgcc gaccccgccc aagccgctca cgtccttcct catccaggac 120
atcctgcggg acggcgcgca gcggcaaggc ggccgcacga gcagccagag acagcgcgac 180
ccggagccgg agccagagcc agagccagag ggaggacgca gccgcgcccgg ggcgcagaac 240

gaccagctga gcaccggggcc ccgcgccgcg ccggatgagg ccgagacgct ggcagagacc 300
gagccagaaa ggcacttgagg gtcttatctg ttggactctg aaaacacttc aggcgccctt 360
ccaaggcttc cccaaacccc taagcagccg cagaagcgct cccgagctgc cttctccac 420
actcaggtga tcgagttgga gaggaagttc agccatcaga agtacctgtc ggcccctgaa 480
cggggccacc tggccaagaa cctcaagctc acggagaccc aagtgaagat atggttccag 540
aacagacgct ataagactaa gcgaaagcag ctctcctcgg agctgggaga cttggagaag 600
cactcctttt tgccggccct gaaagaggag gccttctccc gggcctccct ggtctccgtg 660
tataacagct atccttacta cccatacctg cactgcgtgg gcagctggag cccagctttt 720
tggtaatga 729

<210> 844

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 844

ctactaagcg ctggagtggag ggatcag

27

<210> 845

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 845

catcgagaat tcactactct ctgactagat gtc

33

<210> 846

<211> 161

<212> PRT

<213> Homo sapiens

<400> 846

Met	Gln	His	His	His	His	His	His	Ala	Gly	Val	Arg	Asp	Gln	Gly	Gln
1				5					10				15		
Gly	Ala	Arg	Trp	Pro	His	Thr	Gly	Lys	Arg	Gly	Pro	Leu	Leu	Gln	Gly
			20					25					30		
Leu	Thr	Trp	Ala	Thr	Gly	Gly	His	Cys	Phe	Ser	Ser	Glu	Glu	Ser	Gly
			35				40					45			
Ala	Val	Asp	Gly	Ala	Gly	Gln	Lys	Lys	Asp	Arg	Ala	Trp	Leu	Arg	Cys
			50				55				60				
Pro	Glu	Ala	Val	Ala	Gly	Phe	Pro	Leu	Gly	Ser	Asp	Cys	Arg	Glu	Gly
			65			70				75				80	
Gly	Arg	Gln	Gly	Cys	Gly	Gly	Ser	Asp	Asp	Glu	Asp	Asp	Leu	Gly	Val
			85					90					95		
Ala	Pro	Gly	Leu	Ala	Pro	Ala	Trp	Ala	Leu	Thr	Gln	Pro	Pro	Ser	Gln

	100		105		110										
Ser	Pro	Gly	Pro	Gln	Ser	Leu	Pro	Ser	Thr	Pro	Ser	Ser	Ile	Trp	Pro
	115						120					125			
Gln	Trp	Val	Ile	Leu	Ile	Thr	Glu	Leu	Thr	Ile	Pro	Ser	Pro	Ala	His
	130					135					140				
Gly	Pro	Pro	Trp	Leu	Pro	Asn	Ala	Leu	Glu	Arg	Gly	His	Leu	Val	Arg
145					150					155					160
Glu															

<210> 847

<211> 489

<212> DNA

<213> Homo sapiens

<400> 847

atgcagcatc	accaccatca	ccacgctgga	gtgagggatc	aggggacaggg	cgcgagatgg	60
cctcacacag	ggaagagagg	gcccctcctg	cagggcctca	cctggggccac	aggaggacac	120
tgcttttctt	ctgaggagtc	aggagctgtg	gatggtgctg	gacagaagaa	ggacagggcc	180
tggctcaggt	gtccagaggc	tgctgctggc	ttcccttttg	gatcagactg	cagggagggg	240
gggcggcagg	ggtgtggggg	gagtgcacgat	gaggatgacc	tgggggtggc	tccaggcctt	300
gcccctgcct	gggccctcac	ccagcctccc	tcacagtctc	ctggccctca	gtctctcccc	360
tccactccat	cctccatctg	gcctcagtgg	gtcattctga	tactgaact	gaccataccc	420
agccctgccc	acggccctcc	atggctcccc	aatgccctgg	agaggggaca	tctagtcaga	480
gagtagtga						489

<210> 848

<211> 132

<212> PRT

<213> Homo sapiens

<400> 848

Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu	Ser	Gln	Gly	Gly	Gln	Gly	Phe
1				5					10					15	
Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala	Ile	Ala	Gly	Gln	Ile	Arg	Ser
		20						25					30		
Gly	Gly	Gly	Ser	Pro	Thr	Val	His	Ile	Gly	Pro	Thr	Ala	Phe	Leu	Gly
	35					40					45				
Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn	Gly	Ala	Arg	Val	Gln	Arg	Val
	50				55						60				
Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu	Gly	Ile	Ser	Thr	Gly	Asp	Val
65				70					75					80	
Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile	Asn	Ser	Ala	Thr	Ala	Met	Ala
			85					90						95	
Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly	Asp	Val	Ile	Ser	Val	Asn	Trp
	100							105					110		
Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr	Gly	Asn	Val	Thr	Leu	Ala	Glu
	115						120					125			
Gly	Pro	Pro	Ala												
	130														

009050"624550

```
<210> 852
<211> 400
<212> PRT
```

<213> Homo sapiens

<400> 852

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
      5                      10                      15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
      20                      25                      30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
      35                      40                      45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
      50                      55                      60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
      65                      70                      75                      80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
      85                      90                      95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
      100                     105                     110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
      115                     120                     125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Ile Thr Tyr Val Pro Pro Leu
      130                     135                     140

Leu Leu Glu Val Gly Val Glu Glu Lys Phe Met Thr Met Val Leu Gly
      145                     150                     155                     160

Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala
      165                     170                     175

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
      180                     185                     190

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala
      195                     200                     205

Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu
      210                     215                     220

Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys Gly Gln Val
      225                     230                     235                     240

Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro
      245                     250                     255

Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu
      260                     265                     270

```

000059 "6225350

Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser
275 280 285

Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu
290 295 300

Leu Thr Leu Ile Phe Leu Thr Cys Val Ala Ala Thr Leu Leu Val Ala
305 310 315 320

Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala
325 330 335

Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe
340 345 350

Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg
355 360 365

Met Pro Arg Thr Leu Arg Arg Leu Phe Val Ala Glu Leu Cys Ser Trp
370 375 380

Met Ala Leu Met Thr Phe Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu
385 390 395 400

<210> 853

<211> 20

<212> PRT

<213> Homo sapiens

<400> 853

Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val
5 10 15

Ser Val Arg Val
20

<210> 854

<211> 60

<212> DNA

<213> Homo sapiens

<400> 854

ctgctccac ctccaccgc gctctgcggg gcctctgcct gtgatgtctc cgtacgtgtg 60

<210> 855

<211> 10

<212> PRT

<213> Homo sapiens

<400> 855

Ala Ser Ala Cys Asp Val Ser Val Arg Val
5 10

000000"000000

```
<400> 856
gcctctgcct gtgatgtctc cgtacgtgtg
```

30

```
<400> 857
Ala Ser Ala Cys Asp Val Ser Val Arg
 1             5
```

<400> 858
Ser Ala Cys Asp Val Ser Val Arg Val
5

```
<400> 859
tctgcctgtg atgtctccgt acgtgtg
```

27

<400> 860
Gly Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser
 5 10 15

Ala Ser Asp

```
<210> 861
<211> 19
<212> PRT
<213> Homo sapiens
```

<400> 861

Val Pro Pro Leu Leu Leu Glu Val Gly Val Glu Glu Lys Phe Met Thr
 5 10 15

Met Val Leu

<210> 862

<211> 19

<212> PRT

<213> Homo sapiens

<400> 862

Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
 5 10 15

Gln Leu Leu

<210> 863

<211> 57

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(57)

<223> n = A,T,C or G

<400> 863

ggcnathggnc cngtnytngg nytngtntgy gtnccnytny tnggnwsngc nwsngay 57

<210> 864

<211> 57

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(57)

<223> n = A,T,C or G

<400> 864

gtnccncny tnytnytnga rgtnggngtn gargaraart tyatgacnat ggtnytn 57

<210> 865

<211> 57

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

009050"626360

<223> n = A, T, C or G

atgggtncarm gnytntgggt nwsnmgnytn ytnmgncaym gnaargcnca rytnytn 57

<211> 9

<213> Homo sapiens

Val Leu Gln Cys Val Asn Val Ser Val
1 5

<211> 9

<213> Homo sapiens

Arg Met Pro Thr Val Leu Gln Cys Val
1 5

<211> 9

<213> Homo sapiens

Asn Leu Cys Lys Phe Thr Glu Trp Ile
1 5

<211> 9

<213> Homo sapiens

Met Leu Ile Lys Leu Asp Glu Ser Val
1 5

<211> 9

<213> Homo sapiens

Leu Leu Ala Asn Asp Leu Met Leu Ile
1 5

<210> 871


```
<400> 871
Leu Leu Ala Asn Gly Arg Met Pro Thr Val
  1                      5                      10
```

```
<400> 872
Leu Met Leu Ile Lys Leu Asp Glu Ser Val
  1                      5                      10
```

```
<400> 873
Val Leu Gln Cys Val Asn Val Ser Val Val
  1                      5                10
```

<400> 874
Gly Leu Leu Ala Asn Gly Arg Met Pro Thr
1 5 10

```
<400> 875
Thr Val Leu Gln Cys Val Asn Val Ser Val
  1                5                10
```

<400> 876
Gly Val Leu Val His Pro Gln Trp Val
1 5

<211> 9

<212> PRT

<213> Homo sapiens

<400> 877

Val Leu Val His Pro Gln Trp Val Leu

1

5

[illegible]